

113967

From: Chan, Christina
Sent: Tuesday, February 10, 2004 12:29 PM
To: Davis, Minh-Tam; STIC-Biotech/ChemLib
Subject: RE: Rush search request for 0/9821812

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

CRFF

-----Original Message-----

From: Davis, Minh-Tam
Sent: Tuesday, February 10, 2004 10:05 AM
To: Chan, Christina
Subject: Rush search request for 0/9821812

Please search in commercial database, issued patent file, PGPUB:

- 1) Oligomer search for the amino acid of SEQ ID NO:5.
- 2) Search for SEQ ID NO:5, with size limitation of no more than 50 amino acids.

Thank you.

MINH TAM DAVIS

ART UNIT 1642,

RESEM ROOM 3A24, MB 3C18.

272-0830

Searcher: [Signature]
Phone: 22504
Location: _____
Date Picked Up: 2/10
Date Completed: 2/11
Searcher Prep/Review: _____
Clerical: 10
Online time: 20

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: ✓
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: ✓
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 113967

TO: Minh-Tam Davis
Location: Rem 3a24 / 3c18
Wednesday, February 11, 2004
Art Unit: 1642
Phone: 272-0830
Serial Number: 09 / 821812

From: Jan Delaval
Location: Biotech-Chem Library
Rem 1A51
Phone: 272-2504

jan.delaval@uspto.gov

Search Notes

3/21:
note: Cannot use some of the ref in
here for 102, because of the limitation "ARPs
(for cl 26)
polypeptide", which as shown in the
specification on p. 12 as "androgen regulated
prostate".

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 13:27:46 ; Search time 39 Seconds
(without alignments)
2883.033 Million cell updates/sec

Title: US-09-821-812-5

Perfect score: 537

Sequence: 1 GGLTTTGTGRLGDRPRLS.....PDHVTQHLKQLRQHFAML 537

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 801455 seqs, 209382283 residues

Word size : 0

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	536	99.8	537	12	US-09-821-812-5
2	188	35.0	276	12	US-10-104-047-3850
3	166	30.9	220	9	US-09-925-299-926
4	166	30.9	220	11	US-09-925-299-926
5	7	1.3	75	9	US-09-815-242-10459
6	7	1.3	75	9	US-09-815-242-10459
7	7	1.3	75	9	US-09-815-242-10459
8	7	1.3	75	9	US-09-815-242-10459
9	7	1.3	117	12	US-09-864-408A-8606
10	7	1.3	157	12	US-10-108-260A-2923
11	7	1.3	168	12	US-10-108-260A-2923
12	7	1.3	205	15	US-10-156-761-14485
13	7	1.3	215	11	US-09-847-208-85
14	7	1.3	215	14	US-10-024-955-2
15	7	1.3	267	12	US-10-084-843-142
16	137	1.3	12	12	US-10-193-002-137
17	137	1.3	15	15	US-10-189-346-89
18	137	1.3	15	15	US-10-369-493-17898
19	137	1.3	15	15	US-10-156-761-7888
20	137	1.3	15	15	US-10-369-493-11862
21	137	1.3	15	15	US-10-369-493-14554
22	137	1.3	15	15	US-10-369-493-15128
23	137	1.3	15	15	US-10-369-493-9445
24	137	1.3	15	15	US-10-369-493-17586
25	137	1.3	15	15	US-09-927-827-58
26	137	1.3	15	15	US-10-369-493-21611
27	137	1.3	15	15	US-10-369-493-15491
28	137	1.3	15	15	US-10-369-493-16237
29	137	1.3	15	15	US-10-156-761-9704
30	137	1.3	15	15	US-10-369-493-20653
31	137	1.3	15	15	US-10-369-493-17874
32	137	1.3	15	15	US-10-369-493-733
33	137	1.3	15	15	US-10-369-493-12348
34	137	1.3	15	15	US-09-927-827-62
35	137	1.3	15	15	US-10-369-493-6987
36	137	1.3	15	15	US-10-369-493-17118
37	137	1.3	15	15	US-10-104-047-3730
38	137	1.3	15	15	US-10-156-761-9001
39	137	1.3	15	15	US-10-369-493-16749
40	137	1.3	15	15	US-10-369-493-23376
41	137	1.3	15	15	US-10-094-749-2325
42	137	1.3	15	15	US-10-156-761-12384
43	137	1.3	15	15	US-10-289-762-380
44	137	1.3	15	15	US-09-841-132-503
45	137	1.3	15	15	US-10-156-761-12356
46	137	1.3	15	15	US-10-108-260A-4811
47	137	1.3	15	15	US-10-369-493-10746
48	137	1.3	15	15	US-10-309-437-4
49	137	1.3	15	15	US-10-101-464A-953
50	137	1.3	15	15	US-10-029-217A-4
51	137	1.3	15	15	US-10-295-027-930
52	137	1.3	15	15	US-10-156-761-10081
53	137	1.3	15	15	US-10-144-194A-42
54	137	1.3	15	15	US-09-801-368-172
55	137	1.3	15	15	US-10-369-493-1470
56	137	1.3	15	15	US-10-156-761-10436
57	137	1.3	15	15	US-09-573-822C-139
58	137	1.3	15	15	US-09-941-611-14
59	137	1.3	15	15	US-09-941-611-15
60	137	1.3	15	15	US-10-280-066-369
61	137	1.3	15	15	US-10-044-995-14
62	137	1.3	15	15	US-10-115-072-19
63	137	1.3	15	15	US-09-962-756-541
64	137	1.3	15	15	US-10-371-540-5
65	137	1.3	15	15	US-10-253-471-541
66	137	1.3	15	15	US-10-097-065-585
67	137	1.3	15	15	US-08-424-550B-562
68	137	1.3	15	15	US-09-864-761-38871
69	137	1.3	15	15	US-10-371-540-2
70	137	1.3	15	15	US-09-864-761-41175
71	137	1.3	15	15	US-10-029-386-28750
72	137	1.3	15	15	US-09-864-761-33396
73	137	1.3	15	15	US-10-029-386-30673
74	137	1.3	15	15	US-09-864-761-45530
75	137	1.3	15	15	US-10-029-386-30646
76	137	1.3	15	15	US-09-864-761-35656
77	137	1.3	15	15	US-09-925-302-603
78	137	1.3	15	15	US-09-972-175-33
79	137	1.3	15	15	US-09-972-175-34
80	137	1.3	15	15	US-10-200-522-33
81	137	1.3	15	15	US-10-200-522-34
82	137	1.3	15	15	US-09-864-408A-1798
83	137	1.3	15	15	US-09-864-408A-3108
84	137	1.3	15	15	US-10-050-704-185
85	137	1.3	15	15	US-10-106-698-5574
86	137	1.3	15	15	Sequence 137, App
87	137	1.3	15	15	Sequence 89, Appl
88	137	1.3	15	15	Sequence 17898, A
89	137	1.3	15	15	Sequence 7888, Ap
90	137	1.3	15	15	Sequence 11662, A
91	137	1.3	15	15	Sequence 14554, A
92	137	1.3	15	15	Sequence 15128, A
93	137	1.3	15	15	Sequence 9445, Ap
94	137	1.3	15	15	Sequence 17586, A
95	137	1.3	15	15	Sequence 58, Appl
96	137	1.3	15	15	Sequence 21611, A
97	137	1.3	15	15	Sequence 15491, A
98	137	1.3	15	15	Sequence 16237, A
99	137	1.3	15	15	Sequence 9704, Ap
100	137	1.3	15	15	Sequence 20653, A
101	137	1.3	15	15	Sequence 17874, A
102	137	1.3	15	15	Sequence 733, App
103	137	1.3	15	15	Sequence 12348, A
104	137	1.3	15	15	Sequence 62, Appl
105	137	1.3	15	15	Sequence 6987, Ap
106	137	1.3	15	15	Sequence 17118, A
107	137	1.3	15	15	Sequence 3730, Ap
108	137	1.3	15	15	Sequence 9001, Ap
109	137	1.3	15	15	Sequence 16749, A
110	137	1.3	15	15	Sequence 23376, A
111	137	1.3	15	15	Sequence 2325, Ap
112	137	1.3	15	15	Sequence 12384, A
113	137	1.3	15	15	Sequence 380, App
114	137	1.3	15	15	Sequence 503, App
115	137	1.3	15	15	Sequence 12356, A
116	137	1.3	15	15	Sequence 4811, Ap
117	137	1.3	15	15	Sequence 10746, A
118	137	1.3	15	15	Sequence 4, Appli
119	137	1.3	15	15	Sequence 953, App
120	137	1.3	15	15	Sequence 4, Appli
121	137	1.3	15	15	Sequence 930, App
122	137	1.3	15	15	Sequence 10081, A
123	137	1.3	15	15	Sequence 50, Appl
124	137	1.3	15	15	Sequence 42, Appl
125	137	1.3	15	15	Sequence 172, App
126	137	1.3	15	15	Sequence 1470, Ap
127	137	1.3	15	15	Sequence 10436, A
128	137	1.3	15	15	Sequence 139, App
129	137	1.3	15	15	Sequence 14, Appl
130	137	1.3	15	15	Sequence 15, Appl
131	137	1.3	15	15	Sequence 369, App
132	137	1.3	15	15	Sequence 14, Appl
133	137	1.3	15	15	Sequence 19, Appl
134	137	1.3	15	15	Sequence 541, App
135	137	1.3	15	15	Sequence 5, Appli
136	137	1.3	15	15	Sequence 541, App
137	137	1.3	15	15	Sequence 585, App
138	137	1.3	15	15	Sequence 562, App
139	137	1.3	15	15	Sequence 38871, A
140	137	1.3	15	15	Sequence 2, Appli
141	137	1.3	15	15	Sequence 41175, A
142	137	1.3	15	15	Sequence 28750, A
143	137	1.3	15	15	Sequence 33396, A
144	137	1.3	15	15	Sequence 30673, A
145	137	1.3	15	15	Sequence 45530, A
146	137	1.3	15	15	Sequence 30646, A
147	137	1.3	15	15	Sequence 35656, A
148	137	1.3	15	15	Sequence 603, App
149	137	1.3	15	15	Sequence 34, Appl
150	137	1.3	15	15	Sequence 33, Appl
151	137	1.3	15	15	Sequence 34, Appl
152	137	1.3	15	15	Sequence 1798, Ap
153	137	1.3	15	15	Sequence 3108, Ap
154	137	1.3	15	15	Sequence 47831, A
155	137	1.3	15	15	Sequence 185, App
156	137	1.3	15	15	Sequence 5574, Ap

89 6 1.1 56 12 US-10-029-386-30941 Sequence 30941, A
90 6 1.1 57 12 US-09-864-408A-7814 Sequence 7814, Ap
91 6 1.1 58 10 US-09-989-919-114 Sequence 114, App
92 6 1.1 60 11 US-09-764-891-2756 Sequence 2756, Ap
93 6 1.1 62 9 US-09-925-299-1147 Sequence 1147, Ap
94 6 1.1 62 11 US-09-925-299-1147 Sequence 1147, Ap
95 6 1.1 62 12 US-09-864-408A-3482 Sequence 3482, Ap
96 6 1.1 63 11 US-09-974-879-201 Sequence 201, App
97 6 1.1 63 11 US-09-788-006-109 Sequence 109, App
98 6 1.1 64 11 US-09-305-736-201 Sequence 201, App
99 6 1.1 64 12 US-09-818-683-201 Sequence 201, App
100 6 1.1 65 8 US-08-808-031A-16 Sequence 16, App

ALIGNMENTS

RESULT 1
US-09-821-812-5
; Sequence 5, Application US/09821812
; Publication No. US20030166520A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biayang
; TITLE OF INVENTION: Nucleic Acids
; FILE REFERENCE: P-IS 4373
; CURRENT APPLICATION NUMBER: US/09/821,812
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(537)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-821-812-5

Query Match 99.8%; Score 536; DB 12; Length 537;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGLTTTTCVTRGLGVRPRLSWSAGPSLAAPAAAMSEMEPLLILAWSYFRRRKQQLCADLCT 60
DB 1 GGLTTTTCVTRGLGVRPRLSWSAGPSLAAPAAAMSEMEPLLILAWSYFRRRKQQLCADLCT 60
QY 61 QMLEKSPYDOAWILKARALTETWYIDEIDVDQEGIAEMMLDENALIAQVPRGTSKLPG 120
DB 61 QMLEKSPYDOAWILKARALTETWYIDEIDVDQEGIAEMMLDENALIAQVPRGTSKLPG 120
QY 121 TNOTGSPQAVRPITQAGRPITGFLRPSQSGRPGTMEQAIPTPTAYTARPTITSSGRF 180
DB 121 TNOTGSPQAVRPITQAGRPITGFLRPSQSGRPGTMEQAIPTPTAYTARPTITSSGRF 180
QY 181 VRLGTASMLTSPDGPPIINSLNLTYSQKPKLAKACLSISIFIMKMLRLWILWALSTE 240
DB 181 VRLGTASMLTSPDGPPIINSLNLTYSQKPKLAKACLSISIFIMKMLRLWILWALSTE 240
QY 241 HSQYKDWKWKVQIGKCYRGLGMYRAEKQFSAKQEQEMVDTFLYLKVVYSLDQPVTL 300
DB 241 HSQYKDWKWKVQIGKCYRGLGMYRAEKQFSAKQEQEMVDTFLYLKVVYSLDQPVTL 300
QY 301 NLFKQGLDFPGEVTLTLCGARIYEEMNNSSAAEYKEVLKQDNTHVXAIAICISNHFY 360
DB 301 NLFKQGLDFPGEVTLTLCGARIYEEMNNSSAAEYKEVLKQDNTHVXAIAICISNHFY 360
QY 361 SDQPEIALRYRRLQMGYINGOLFNNLGLCCFYAQOYDMTLTSPERALSLAENEEAAD 420
DB 361 SDQPEIALRYRRLQMGYINGOLFNNLGLCCFYAQOYDMTLTSPERALSLAENEEAAD 420
QY 421 VVYNLGHVAVGIGDTNLAHQCFRLALVNNNNHAEAYNNLAVLEMRKGHVQEARALLQTAS 480

DB 421 VVYNLGHVAVGIGDTNLAHQCFRLALVNNNNHAEAYNNLAVLEMRKGHVQEARALLQTAS 480
QY 481 SLAPHMVEPHNFATISDKIGDLQORSYVAAQKSEAAFPDHDVDTQHLIKQLRQHFAML 537
DB 481 SLAPHMVEPHNFATISDKIGDLQORSYVAAQKSEAAFPDHDVDTQHLIKQLRQHFAML 537
RESULT 2
US-10-104-047-3850
; Sequence 3850, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1 full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3850
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3850
Query Match 35.0%; Score 188; DB 12; Length 276;
Best Local Similarity 100.0%; Pred. No. 3.6e-178;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 350 AIACIGSNHFYSQDPETALRYRRLQMGYINGOLFNNLGLCCFYAQOYDMTLTSPERAL 409
DB 89 AIACIGSNHFYSQDPETALRYRRLQMGYINGOLFNNLGLCCFYAQOYDMTLTSPERAL 148
QY 410 SLAENEEAADVWNLGHVAVGIGDTNLAHQCFRLALVNNNNHAEAYNNLAVLEMRKGHV 469
DB 149 SLAENEEAADVWNLGHVAVGIGDTNLAHQCFRLALVNNNNHAEAYNNLAVLEMRKGHV 208
QY 470 EQARALLQTASSLAPHMVEPHNFATISDKIGDLQORSYVAAQKSEAAFPDHDVDTQHLIKQ 529
DB 209 EQARALLQTASSLAPHMVEPHNFATISDKIGDLQORSYVAAQKSEAAFPDHDVDTQHLIKQ 268
QY 530 LRQHFAML 537
DB 269 LRQHFAML 276

RESULT 3
US-09-925-299-926
; Sequence 926, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 926
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-926
Query Match 30.9%; Score 166; DB 9; Length 220;
Best Local Similarity 100.0%; Pred. No. 2.2e-156;

Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 372 RRLQMGYNGQLFNNLGLCCFYAQYDMTILTSFERALSLAENBEEAADVWYNLGHVAVG 431
Db 55 RRLQMGYNGQLFNNLGLCCFYAQYDMTILTSFERALSLAENBEEAADVWYNLGHVAVG 114

Qy 432 IGDNLAHQCFLALVNNNNHAEAYNNLAVLEMRKGHVEQARALLQTASSLAPHMYEPHF 491
Db 115 IGDNLAHQCFLALVNNNNHAEAYNNLAVLEMRKGHVEQARALLQTASSLAPHMYEPHF 174

Qy 492 NFATISDKIGLQRSYVAQKSEAAFPDHDVDTQHLIKOLRQHFAML 537
Db 175 NFATISDKIGLQRSYVAQKSEAAFPDHDVDTQHLIKOLRQHFAML 220

RESULT 4

US-09-925-299-926
; Sequence 926, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 926
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-926

Query Match 30.9%; Score 166; DB 11; Length 220;
Best Local Similarity 100.0%; Pred. No. 2.26-156;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 372 RRLQMGYNGQLFNNLGLCCFYAQYDMTILTSFERALSLAENBEEAADVWYNLGHVAVG 431
Db 55 RRLQMGYNGQLFNNLGLCCFYAQYDMTILTSFERALSLAENBEEAADVWYNLGHVAVG 114

Qy 432 IGDNLAHQCFLALVNNNNHAEAYNNLAVLEMRKGHVEQARALLQTASSLAPHMYEPHF 491
Db 115 IGDNLAHQCFLALVNNNNHAEAYNNLAVLEMRKGHVEQARALLQTASSLAPHMYEPHF 174

Qy 492 NFATISDKIGLQRSYVAQKSEAAFPDHDVDTQHLIKOLRQHFAML 537
Db 175 NFATISDKIGLQRSYVAQKSEAAFPDHDVDTQHLIKOLRQHFAML 220

RESULT 5

US-09-815-242-10459
; Sequence 10459, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10459
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10459

Query Match 1.3%; Score 7; DB 9; Length 75;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 YFRRKRF 52
Db 4 YFRRKRF 10

RESULT 6

US-09-815-242-11050
; Sequence 11050, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11050
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11050

Query Match 1.3%; Score 7; DB 9; Length 75;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 YFRRRK 52
|||||
Db 4 YFRRRK 10

RESULT 7

US-09-815-242-13859
; Sequence 13859, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13859
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-13859

Query Match 1.3%; Score 7; DB 9; Length 75;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 YFRRRK 52
|||||
Db 4 YFRRRK 10

RESULT 8

US-09-864-408A-8606
; Sequence 8606, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8606
; LENGTH: 85
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-864-408A-8606

Query Match 1.3%; Score 7; DB 12; Length 85;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 QSGRPGT 156
|||||
Db 20 QSGRPGT 26

RESULT 9

US-10-029-386-34055
; Sequence 34055, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: A601CA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34055
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004595.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.86
; OTHER INFORMATION: SWISSPROT HIT: P23098, EVALUATE 2.60e-02
US-10-029-386-34055

Query Match 1.3%; Score 7; DB 12; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 ENEEEAA 419
|||||
Db 40 ENEEEAA 46

RESULT 10

US-10-108-260A-2923
; Sequence 2923, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2923
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-2923

Query Match 1.3%; Score 7; DB 12; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 ENEEEAA 419
|||||
Db 40 ENEEEAA 46

```
RESULT 11
US-10-094-749-2002
; Sequence 2002, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2002
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2002

Query Match          1.3%; Score 7; DB 12; Length 168;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 QARALLQ 477
Db 9 QARALLQ 15

RESULT 12
US-10-156-761-14485
; Sequence 14485, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14485
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
```

```
US-10-156-761-14485

Query Match          1.3%; Score 7; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 510 AQKSEAA 516
Db 40 AQKSEAA 46

RESULT 13
US-09-847-208-85
; Sequence 85, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 85
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus (House-dust mite)
US-09-847-208-85

Query Match          1.3%; Score 7; DB 11; Length 215;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 400 MTLTSFE 406
Db 152 MTLTSFE 158

RESULT 14
US-10-024-955-2
; Sequence 2, Application US/10024955
; Publication No. US20020168373A1
; GENERAL INFORMATION:
; APPLICANT: Wayne R. Thomas and Kaw-Yan Chua
; TITLE OF INVENTION: Allergenic Proteins and Peptides From
; TITLE OF INVENTION: House Dust Mite and Uses Therefor
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/024,955
; FILING DATE: 19-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,336A
; FILING DATE: 10-JUN-1996
; APPLICATION NUMBER: US 08/081,540
; FILING DATE: 22-JUNE-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
```

REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: IMI-032CP2
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 215 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-024-955-2

Query Match 1.3%; Score 7; DB 14; Length 215;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 400 MTLTSFE 406
Db 152 MTLTSFE 158

RESULT 15
US-10-084-843-142
; Sequence 142, Application US/10084843
; Publication No. US20030143243A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; Skeiky, Yasir A.W.
; Dillon, Davin C.
; Campos-Neto, Antonio
; Houghton, Raymond
; Vedvick, Thomas S.
; Twardzik, Daniel R.
; Lodes, Michael J.
; Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,843
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 142:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-10-084-843-142

Query Match 1.3%; Score 7; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 VPRPGTS 115
Db 78 VPRPGTS 84

RESULT 16
US-10-193-002-137
; Sequence 137, Application US/10193002
; Publication No. US20030135026A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; Skeiky, Yasir A.W.
; Dillon, Davin C.
; Campos-Neto, Antonio
; Houghton, Raymond
; Vedvick, Thomas S.
; Twardzik, Daniel R.
; Lodes, Michael J.
; Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/193,002
; FILING DATE: 10-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 137:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 137:
US-10-193-002-137

Query Match 1.3%; Score 7; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 VPRPGTS 115
Db 78 VPRPGTS 84

RESULT 17
US-10-189-346-89
; Sequence 89, Application US/10189346
; Publication No. US20030119162A1
; GENERAL INFORMATION:
; APPLICANT: Churchill, Mair
; APPLICANT: von Bodman, Susanne
; APPLICANT: Schweizer, Herbert
; APPLICANT: Gould, Ty
; APPLICANT: Hoang, Tung
; APPLICANT: Murphy, Frank
; APPLICANT: Watson, William
; TITLE OF INVENTION: Structural Basis of Quorum Sensing Signal Generation and Methods
; FILE REFERENCE: 2848-44
; CURRENT APPLICATION NUMBER: US/10/189,346
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303,449
; PRIOR FILING DATE: 2001-07-04
; PRIOR APPLICATION NUMBER: 60/366,575
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
; LENGTH: 267
; TYPE: PRT
; ORGANISM: No. US20030119162Altoc BP.
US-10-189-346-89

Query Match 1.3%; Score 7; DB 15; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 411 LAENEE 417
|||||||
Db 33 LAENEE 39

RESULT 18
US-10-369-493-17898
; Sequence 17898, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17898
; LENGTH: 291
; TYPE: PRT
; ORGANISM: SPHINGOMONAS
US-10-369-493-17898

Query Match 1.3%; Score 7; DB 12; Length 291;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 74 ILKARAL 80
|||||||
Db 103 ILKARAL 109

RESULT 19

US-10-156-761-7888
; Sequence 7888, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: HORIKAWA, JUN
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7888
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7888

Query Match 1.3%; Score 7; DB 15; Length 303;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 469 VEQARAL 475
|||||||
Db 36 VEQARAL 42

RESULT 20
US-10-369-493-11662
; Sequence 11662, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11662
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-11662

Query Match 1.3%; Score 7; DB 12; Length 307;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 9 GTRLGVD 15
|||||||
Db 17 GTRLGVD 23

RESULT 21
US-10-369-493-14554
; Sequence 14554, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

```
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14554
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14554

Query Match      1.3%; Score 7; DB 12; Length 307;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      9 GTRLGVD 15
Db      17 GTRLGVD 23

RESULT 22
US-10-369-493-15128
; Sequence 15128, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15128
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-15128

Query Match      1.3%; Score 7; DB 12; Length 307;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      9 GTRLGVD 15
Db      17 GTRLGVD 23

RESULT 23
US-10-369-493-9445
; Sequence 9445, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
```

```
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9445
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-369-493-9445

Query Match      1.3%; Score 7; DB 12; Length 335;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      470 EQARALL 476
Db      257 EQARALL 263

RESULT 24
US-10-369-493-17586
; Sequence 17586, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17586
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-369-493-17586

Query Match      1.3%; Score 7; DB 12; Length 344;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      470 EQARALL 476
Db      264 EQARALL 270

RESULT 25
US-09-927-827-58
; Sequence 58, Application US/09927827
; Publication No. US20030036176A1
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; APPLICANT: Ramseier, Thomas M.
; TITLE OF INVENTION: Directed Genetic Engineering of Xanthomonas campestris
; FILE REFERENCE: 38-10(15824)B
; CURRENT APPLICATION NUMBER: US/09/927,827
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/279,493
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 69
; SEQ ID NO 58
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-09-927-827-58
```


Query Match 1.3%; Score 7; DB 11; Length 348;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 428 VAVGIGD 434
Db 13 VAVGIGD 19

RESULT 26
US-10-369-493-21611
; Sequence 21611, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21611
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Pyrococcus abyssi
US-10-369-493-21611

Query Match 1.3%; Score 7; DB 12; Length 363;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 405 FERALSL 411
Db 16 FERALSL 22

RESULT 27
US-10-369-493-15491
; Sequence 15491, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15491
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15491

Query Match 1.3%; Score 7; DB 12; Length 388;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 406 ERALSIA 412
Db 16 ERALSIA 22

Db 7 ERALSIA 13

RESULT 28
US-10-369-493-16237
; Sequence 16237, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16237
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-16237

Query Match 1.3%; Score 7; DB 12; Length 388;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 406 ERALSIA 412
Db 7 ERALSIA 13

RESULT 29
US-10-156-761-9704
; Sequence 9704, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9704
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9704

Query Match 1.3%; Score 7; DB 15; Length 404;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 470 EQBALL 476
Db 147 EQBALL 153

RESULT 30

```
US-10-369-493-20653
; Sequence 20653, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20653
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Rhodospseudomonas palustris
US-10-369-493-20653

Query Match 1.3%; Score 7; DB 12; Length 418;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VYIDEID 90
DB 180 VYIDEID 186

RESULT 31
US-10-369-493-17874
; Sequence 17874, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17874
; LENGTH: 421
; TYPE: PRT
; ORGANISM: SPHINGOMONAS
US-10-369-493-17874

Query Match 1.3%; Score 7; DB 12; Length 421;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VYIDEID 90
DB 181 VYIDEID 187

RESULT 32
US-10-369-493-733
; Sequence 733, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 733
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-733

Query Match 1.3%; Score 7; DB 12; Length 424;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VYIDEID 90
DB 181 VYIDEID 187

RESULT 33
US-10-369-493-12348
; Sequence 12348, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12348
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-12348

Query Match 1.3%; Score 7; DB 12; Length 424;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VYIDEID 90
DB 182 VYIDEID 188

RESULT 34
US-09-927-827-62
; Sequence 62, Application US/09927827
; Publication No. US20030036176A1
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; APPLICANT: Ramseier, Thomas M.
; TITLE OF INVENTION: Directed Genetic Engineering of Xanthomonas campestris
; FILE REFERENCE: 38-10(15824)B
; CURRENT APPLICATION NUMBER: US/09/927,827
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/279,493
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 69
; SEQ ID NO 62
```

```
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-09-827-827-62

Query Match      1.3%; Score 7; DB 11; Length 428;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      84 VYIDEID 90
Db      183 VYIDEID 189

RESULT 35
US-10-369-493-6987
; Sequence 6987, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6987
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6987

Query Match      1.3%; Score 7; DB 12; Length 437;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      177 SGRFVRL 183
Db      398 SGRFVRL 404

RESULT 36
US-10-369-493-17118
; Sequence 17118, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17118
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Bacillus halodurans
US-10-369-493-17118

Query Match      1.3%; Score 7; DB 12; Length 447;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
```

```
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      479 ASSLAPH 485
Db      184 ASSLAPH 190

RESULT 37
US-10-104-047-3730
; Sequence 3730, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3730
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3730

Query Match      1.3%; Score 7; DB 12; Length 470;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      112 PGTSLKL 118
Db      33 PGTSLKL 39

RESULT 38
US-10-156-761-9001
; Sequence 9001, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9001
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9001

Query Match      1.3%; Score 7; DB 15; Length 493;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      183 LGTASML 189
Db      214 LGTASML 220

RESULT 39
```

```
US-10-369-493-16749
; Sequence 16749, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16749
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-369-493-16749

Query Match          1.3%; Score 7; DB 12; Length 513;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 FINLSRL 202
DB 488 FINLSRL 494

RESULT 40
US-10-369-493-23376
; Sequence 23376, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23376
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-369-493-23376

Query Match          1.3%; Score 7; DB 12; Length 513;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 FINLSRL 202
DB 488 FINLSRL 494

RESULT 41
US-10-094-749-2325
; Sequence 2325, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
```

```
US-10-156-761-12384
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTYUUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2325
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2325

Query Match          1.3%; Score 7; DB 12; Length 524;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 KARALTE 82
DB 442 KARALTE 448

RESULT 42
US-10-156-761-12384
; Sequence 12384, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12384
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12384

Query Match          1.3%; Score 7; DB 15; Length 542;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LTTTVIG 9
|||||
```

Db 515 LTTTVIG 521

RESULT 43

US-10-289-762-380

Sequence 380, Application US/10289762

Publication No. US20040006218A1

GENERAL INFORMATION:

APPLICANT: Griffais, R.

TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection

TITLE OF INVENTION: and treatment of infection

FILE REFERENCE: 9710-003-999

CURRENT APPLICATION NUMBER: US/10/289,762

PRIOR FILING DATE: 2003-03-27

NUMBER OF SEQ ID NOS: 6849

SEQ ID NO 380

LENGTH: 544

TYPE: PRT

ORGANISM: Chlamydia pneumoniae

FEATURE:

NAME/KEY: SITE

LOCATION: 1...544

OTHER INFORMATION: Xaa=unknown or other

US-10-289-762-380

Query Match 1.3%; Score 7; DB 12; Length 544;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 528 KQLRQHF 534

Db 225 KQLRQHF 231

RESULT 44

US-09-841-132-503

Sequence 503, Application US/09841132

Patent No. US20020061848A1

GENERAL INFORMATION:

APPLICANT: Bhatia, Ajay

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Probst, Peter

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND DIAGNOSIS OF CHLAMYDIAL INFECTION

TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

FILE REFERENCE: 210121.469C8

CURRENT APPLICATION NUMBER: US/09/841,132

CURRENT FILING DATE: 2001-04-23

NUMBER OF SEQ ID NOS: 599

SOFTWARE: FastSeq for Windows Version 3.0/4.0

SEQ ID NO 503

LENGTH: 582

TYPE: PRT

ORGANISM: Chlamydia pneumoniae

US-09-841-132-503

Query Match 1.3%; Score 7; DB 9; Length 582;

Best Local Similarity 100.0%; Pred. No. 5.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 528 KQLRQHF 534

Db 263 KQLRQHF 269

RESULT 45

US-10-156-761-12356

Sequence 12356, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 12356

LENGTH: 583

TYPE: PRT

ORGANISM: Streptomyces avermitilis

US-10-156-761-12356

Query Match 1.3%; Score 7; DB 15; Length 583;

Best Local Similarity 100.0%; Pred. No. 5.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 AGPSLAA 29

Db 95 AGPSLAA 101

RESULT 46

US-10-108-260A-4811

Sequence 4811, Application US/10108260A

Publication No. US20040005560A1

GENERAL INFORMATION:

APPLICANT: HELIX RESEARCH INSTITUTE

TITLE OF INVENTION: No. US20040005560A1 full length cDNA

FILE REFERENCE: HI-A0106

CURRENT APPLICATION NUMBER: US/10/108,260A

CURRENT FILING DATE: 2002-03-27

NUMBER OF SEQ ID NOS: 5458

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 4811

LENGTH: 587

TYPE: PRT

ORGANISM: Homo sapiens

US-10-108-260A-4811

Query Match 1.3%; Score 7; DB 12; Length 587;

Best Local Similarity 100.0%; Pred. No. 5.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 337 YKEVLKQ 343

Db 118 YKEVLKQ 124

RESULT 47

US-10-369-493-10746

Sequence 10746, Application US/10369493

Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

```
; SEQ ID NO 10746
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Spingomonas aromaticivorans
US-10-369-493-10746

Query Match      1.3%; Score 7; DB 12; Length 622;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      470 EQARALL 476
      |||||
Db      384 EQARALL 390

RESULT 48
US-10-309-437-4
; Sequence 4, Application US/10309437
; Publication No. US20030119164A1
; GENERAL INFORMATION:
; APPLICANT: Udagawa, Hiroaki
; APPLICANT: Frandsen, Torben
; APPLICANT: Nielsen, Tom
; APPLICANT: Kauppinen, Markus
; APPLICANT: Christensen, Soeren
; TITLE OF INVENTION: Lysophospholipase
; FILE REFERENCE: 5958.210-US
; CURRENT APPLICATION NUMBER: US/10/309,437
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: US/09/687,538
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-10-309-437-4

Query Match      1.3%; Score 7; DB 15; Length 638;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      199 LSRNLMT 205
      |||||
Db      90 LSRNLMT 96

RESULT 49
US-10-101-464A-953
; Sequence 953, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: Fastseq for Windows Version 4.0
; SEQ ID NO 953
; LENGTH: 653

; TYPE: PRT
; ORGANISM: Eucaalyptus grandis
US-10-101-464A-953

Query Match      1.3%; Score 7; DB 15; Length 653;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      274 LKQEMV 280
      |||||
Db      284 LKQEMV 290

RESULT 50
US-10-029-217A-4
; Sequence 4, Application US/10029217A
; Publication No. US20020164735A1
; GENERAL INFORMATION:
; APPLICANT: OLSON, ERIC N.
; APPLICANT: WANG, DA-ZHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO A CARDIAC-SPECIFIC
; TITLE OF INVENTION: NUCLEAR REGULATORY FACTOR
; FILE REFERENCE: UTS:695US
; CURRENT APPLICATION NUMBER: US/10/029,217A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/257,761
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-029-217A-4

Query Match      1.3%; Score 7; DB 14; Length 693;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      25 PSLAAPA 31
      |||||
Db      601 PSLAAPA 607

RESULT 51
US-10-295-027-930
; Sequence 930, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
```

```
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 930
; LENGTH: 790
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-930

Query Match          1.3%; Score 7; DB 12; Length 790;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      432 IGDYNLA 438
      |||||||
Db      264 IGDYNLA 270

RESULT 52
US-10-156-761-10081
; Sequence 10081, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10081
; LENGTH: 810
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10081

Query Match          1.3%; Score 7; DB 15; Length 810;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      110 PRPGTSL 116
      |||||||
Db      112 PRPGTSL 118

RESULT 53
US-10-001-873-50
; Sequence 50, Application US/10001873
; Publication No. US20020160388A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
```

```
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0275
; CURRENT APPLICATION NUMBER: US/10/001,873
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,055
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/252,496
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 1134
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-873-50

Query Match          1.3%; Score 7; DB 14; Length 1134;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      24 GPSLAAP 30
      |||||||
Db      818 GPSLAAP 824

RESULT 54
US-10-144-194A-42
; Sequence 42, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-194A-42

Query Match          1.3%; Score 7; DB 12; Length 1179;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      457 NNLAIVL 463
      |||||||
Db      133 NNLAIVL 139

RESULT 55
US-09-801-368-172
; Sequence 172, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
```

; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi

; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 172
; TYPE: PRT
; LENGTH: 3092
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-172

Query Match 1.3%; Score 7; DB 10; Length 3092;

Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LFNNLGL 390
|||||
DB 1402 LFNNLGL 1408

RESULT 56

US-10-369-493-1470
; Sequence 1470, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1470
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1470

Query Match 1.3%; Score 7; DB 12; Length 3092;

Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LFNNLGL 390
|||||
DB 1402 LFNNLGL 1408

RESULT 57

US-10-156-761-10436
; Sequence 10436, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10436
; LENGTH: 6146
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10436

Query Match 1.3%; Score 7; DB 15; Length 6146;

Best Local Similarity 100.0%; Pred. No. 4.7e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 LTSPDGP 195
|||||
DB 3460 LTSPDGP 3466

RESULT 58

US-09-573-822C-139
; Sequence 139, Application US/09573822C
; Publication No. US20030199011A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands generated from microbial genome seq
; FILE REFERENCE: Microbe patent
; CURRENT APPLICATION NUMBER: US/09/573,822C
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 804
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 139
; LENGTH: 10
; TYPE: PRT
; ORGANISM: mycoplasma genitalium
; FEATURE:
; OTHER INFORMATION: Sequence located in MG079 at 250-259 and may interact with Sequen
; OTHER INFORMATION: in this patent.
US-09-573-822C-139

Query Match 1.1%; Score 6; DB 12; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 526 LIKQLR 531
|||||
DB 1 LIKQLR 6

RESULT 59

US-09-941-611-14
; Sequence 14, Application US/09941611
; Patent No. US20020106640A1
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; POLLET, DIRK
; MAERTENS, GEERT
; VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible


```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/941,611
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/391,671
; FILING DATE: 1995-02-21
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-941-611-14

Query Match      1.1%; Score 6; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      475 LLOTAS 480
Db      12 LLOTAS 17
      |||||

RESULT 60
US-09-941-611-15
; Sequence 15, Application US/09941611
; Patent No. US20020106640A1
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; POLLET, DIRK
; MAERTENS, GEERT
; VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/941,611
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/391,671
; FILING DATE: 1995-02-21
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; APPLICATION NUMBER: EP 90124241.2
```

```
;
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-941-611-15

Query Match      1.1%; Score 6; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      475 LLOTAS 480
Db      6 LLOTAS 11
      |||||

RESULT 61
US-10-280-066-369
; Sequence 369, Application US/10280066
; Publication No. US20030180718A1
; GENERAL INFORMATION:
; APPLICANT: Pillutla, Renuka C.
; APPLICANT: Brissette, Renee
; APPLICANT: Spruyt, Michael
; APPLICANT: Dedova, Olga
; APPLICANT: Blume, Arthur J.
; APPLICANT: Prendergast, John
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: TARGET SPECIFIC SCREENING AND ITS USE FOR IDENTIFYING TARGET BIND
; FILE REFERENCE: 2598-4009US1
; CURRENT APPLICATION NUMBER: US/10/280,066
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 60/345,471
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 537
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 369
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Eschericia coli
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: VEGF-20F-3-A1
US-10-280-066-369

Query Match      1.1%; Score 6; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      471 QARALL 476
Db      6 QARALL 11
      |||||

RESULT 62
US-10-044-995-14
; Sequence 14, Application US/10044995
; Publication No. US20030049685A1
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; POLLET, DIRK
; MAERTENS, GEERT
```

```
;
;
; VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; ANTIBODIES TO HEPATITIS C VIRUS
;
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/044,995
; FILING DATE: 15-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/391,671
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/920,286
; FILING DATE: 14-OCT-1992
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
;
; US-10-044-995-14
;
; Query Match 1.1%; Score 6; DB 15; Length 20;
; Best Local Similarity 100.0%; Pred. No. 2.8e+02;
; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 475 LLOQTAS 480
; Db 12 LLOQTAS 17
;
; RESULT 63
; US-10-044-995-15
; Sequence 15, Application US/10044995
; Publication No. US20030049685A1
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; POLLET, DIRK
; MAERTENS, GEERT
; VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; ANTIBODIES TO HEPATITIS C VIRUS
;
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
```

```
;
;
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/044,995
; FILING DATE: 15-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/391,671
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/920,286
; FILING DATE: 14-OCT-1992
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
;
; US-10-044-995-15
;
; Query Match 1.1%; Score 6; DB 15; Length 20;
; Best Local Similarity 100.0%; Pred. No. 2.8e+02;
; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 475 LLOQTAS 480
; Db 6 LLOQTAS 11
;
; RESULT 64
; US-10-115-072-19
; Sequence 19, Application US/10115072
; Publication No. US20030105003A1
; GENERAL INFORMATION:
; APPLICANT: NILSSON, JAN
; APPLICANT: SHAH, PREDIMAN K.
; TITLE OF INVENTION: PEPTIDE-BASED IMMUNIZATION THERAPY FOR TREATMENT OF
; ATHEROSCLEROSIS AND DEVELOPMENT OF PEPTIDE-BASED ASSAY
; TITLE OF INVENTION: FOR DETERMINATION OF IMMUNE RESPONSES AGAINST OXIDIZED
; TITLE OF INVENTION: LOW DENSITY LIPOPROTEIN
; FILE REFERENCE: 03940.0057
; CURRENT APPLICATION NUMBER: US/10/115,072
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/281,410
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: SE 0101232-7
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: SE 0103754-8
; PRIOR FILING DATE: 2001-09-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
```

; OTHER INFORMATION: peptide
US-10-115-072-19

Query Match 1.1%; Score 6; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 204 LTKYSQ 209
Db 5 LTKYSQ 10

RESULT 65

US-09-962-756-541
; Sequence 541, Application US/09962756
; Publication No. US20030195147A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA
; APPLICANT: BRISSETTE, RENEE
; APPLICANT: BLUME, ARTHUR J.
; APPLICANT: SCHAFFER, LAUGE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-40510S1
; CURRENT APPLICATION NUMBER: US/09/962,756
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 541
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-962-756-541

Query Match 1.1%; Score 6; DB 12; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 405 FERALS 410
Db 15 FERALS 20

RESULT 66

US-10-371-540-5
; Sequence 5, Application US/10371540
; Publication No. US20030198644A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; TITLE OF INVENTION: HCV PEPTIDE ANTIGENS AND METHODS FOR THE DETERMINATION OF HCV
; FILE REFERENCE: 9793/129
; CURRENT APPLICATION NUMBER: US/10/371,540
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: 09/689,678
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 08/604,365
; PRIOR FILING DATE: 1996-02-21
; PRIOR APPLICATION NUMBER: P 41 41 304.1
; PRIOR FILING DATE: 1991-12-14
; PRIOR APPLICATION NUMBER: P 42 09 215.9
; PRIOR FILING DATE: 1992-03-21
; PRIOR APPLICATION NUMBER: PCT/EP92/01468

; PRIOR FILING DATE: 1992-06-30
; PRIOR APPLICATION NUMBER: P 41 22 160.5
; PRIOR FILING DATE: 1991-07-04
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: partial sequence of peptide antigen SEQ ID NO. 2 for anti-HCV ant

US-10-371-540-5

Query Match 1.1%; Score 6; DB 12; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLOTAS 480
Db 14 LLOTAS 19

RESULT 67

US-10-253-471-541
; Sequence 541, Application US/10253471
; Publication No. US20030236190A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4057
; CURRENT APPLICATION NUMBER: US/10/253,471
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 541
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-253-471-541

Query Match 1.1%; Score 6; DB 12; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 405 FERALS 410
Db 15 FERALS 20

RESULT 68

US-10-097-065-585
; Sequence 585, Application US/10097065
; Publication No. US20030055236A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021P1
; CURRENT APPLICATION NUMBER: US/10/097,065
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: PCT/US98/27059
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18

;; PRIOR APPLICATION NUMBER: 60/068,007
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,057
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,006
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,369
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,367
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,368
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,169
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,053
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,064
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,054
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,008
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,365
;; PRIOR FILING DATE: 1997-12-19
;; NUMBER OF SEQ ID NOS: 672
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 585
;; LENGTH: 22
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-097-065-585

Query Match 1.1%; Score 6; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 AVGIGD 434
|||
Db 13 AVGIGD 18

RESULT 69
US-08-424-550B-562
; Sequence 562, Application US/08424550B
; Publication No. US2002011947A1
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BULJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550B

;; FILING DATE:
;; CLASSIFICATION: 435435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOREMSKI, PRISCILLA E.
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 5527.PC.01
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-938-2623
;; INFORMATION FOR SEQ ID NO: 562:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 29 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-424-550B-562

Query Match 1.1%; Score 6; DB 8; Length 29;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 LLQTAS 480
|||||
Db 15 LLQTAS 20

RESULT 70
US-09-864-761-38871
; Sequence 38871, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/532,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/006666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006659
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408

;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 38871
;; LENGTH: 32
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO Z98884.11
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
;; OTHER INFORMATION: EST_HUMAN HIT: BE885442.1, EVALUATE 1.50e+00
US-09-864-761-38871

Query Match 1.1%; Score 6; DB 9; Length 32;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 503 LORSYV 508
|||
Db 25 LORSYV 30

RESULT 71

US-10-371-540-2
;; Sequence 2, Application US/10371540
;; Publication No. US20030198644A1
;; GENERAL INFORMATION:
;; APPLICANT: Roche Diagnostics GmbH
;; TITLE OF INVENTION: HCV PEPTIDE ANTIGENS AND METHODS FOR THE DETERMINATION OF HCV
;; FILE REFERENCE: 9793/129
;; CURRENT APPLICATION NUMBER: US/10/371,540
;; CURRENT FILING DATE: 2003-02-21
;; PRIOR APPLICATION NUMBER: 09/689,678
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 08/604,365
;; PRIOR FILING DATE: 1996-02-21
;; PRIOR APPLICATION NUMBER: P 41 41 304.1
;; PRIOR FILING DATE: 1991-12-14
;; PRIOR APPLICATION NUMBER: P 42 09 215.9
;; PRIOR FILING DATE: 1992-03-21
;; PRIOR APPLICATION NUMBER: PCT/EP92/01468
;; PRIOR FILING DATE: 1992-06-30
;; PRIOR APPLICATION NUMBER: P 41 22 160.5
;; PRIOR FILING DATE: 1991-07-04
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 2
;; LENGTH: 33
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: peptide antigen for anti-HCV antibodies
US-10-371-540-2

Query Match 1.1%; Score 6; DB 12; Length 33;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLQTAS 480
|||
Db 26 LLQTAS 31

RESULT 72

US-09-864-761-41175
;; Sequence 41175, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aeomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 41175
;; LENGTH: 37
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC019221.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.5
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.7
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.3
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.3
;; OTHER INFORMATION: EST_HUMAN HIT: AA78178.1, EVALUATE 5.00e-03
US-09-864-761-41175

Query Match 1.1%; Score 6; DB 9; Length 37;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 465 RKGHVE 470

```
Db      8 RKGHVE 13
|||||
RESULT 73
US-10-029-386-28750
; Sequence 28750, Application US/10029386
; Publication No. US200301947041
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28750
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.6
; OTHER INFORMATION: SWISSPROT HIT: P47838, EVALUAE 5.30e-02
US-10-029-386-28750

Query Match      1.1%; Score 6; DB 12; Length 38;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      414 NEEEA 419
|||||
Db      9 NEEEA 14

RESULT 74
US-09-864-761-33396
; Sequence 33396, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/532,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33396
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC00482.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.82
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.7
; OTHER INFORMATION: SWISSPROT HIT: P46662, EVALUAE 4.00e-05
; OTHER INFORMATION: EST_HUMAN HIT: H55159.1, EVALUAE 4.00e-04
US-09-864-761-33396

Query Match      1.1%; Score 6; DB 9; Length 40;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      372 RRLQOM 377
|||||
Db      24 RRLQOM 29

RESULT 75
US-10-029-386-30673
; Sequence 30673, Application US/10029386
; Publication No. US200301947041
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30673
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR11.1
```

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.73
US-10-029-386-30673

Query Match 1.1%; Score 6; DB 12; Length 41;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 YIDEID 90
|||
Db 11 YIDEID 16

Search completed: February 10, 2004, 13:31:01
Job time : 42 secs

THIS PAGE BLANK (user)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 13:26:36 ; Search time 21 Seconds

(without alignment)
1081.949 Million cell updates/sec

Title: US-09-821-812-5

Perfect score: 537

Sequence: 1 GGLTTTIGRLGVDRPLS.....PDHVDTOHLIKQLRQHPAML 537

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents AA.*

1: /cgm2_6/ptodata/1/iaa/5A-COMB.pep.*

2: /cgm2_6/ptodata/1/iaa/5B-COMB.pep.*

3: /cgm2_6/ptodata/1/iaa/6A-COMB.pep.*

4: /cgm2_6/ptodata/1/iaa/6B-COMB.pep.*

5: /cgm2_6/ptodata/1/iaa/PCTUS-COMB.pep.*

6: /cgm2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	1.5	123	4	US-09-462-606-63
2	7	1.3	29	1	Sequence 63, Appl
3	7	1.3	74	4	Sequence 1, Appl
4	7	1.3	74	4	Sequence 1483, Ap
5	7	1.3	74	4	Sequence 1485, Ap
6	7	1.3	140	4	Sequence 1678, Ap
7	7	1.3	198	4	Sequence 18398, A
8	7	1.3	210	4	Sequence 18398, A
9	7	1.3	215	3	Sequence 2, Appl
10	7	1.3	215	4	Sequence 2, Appl
11	7	1.3	244	4	Sequence 5529, Ap
12	7	1.3	267	3	Sequence 142, App
13	7	1.3	267	4	Sequence 137, App
14	7	1.3	267	4	Sequence 142, App
15	7	1.3	302	4	Sequence 173, App
16	7	1.3	304	4	Sequence 173, App
17	7	1.3	306	4	Sequence 6459, Ap
18	7	1.3	315	3	Sequence 17986, A
19	7	1.3	327	4	Sequence 4, Appl
20	7	1.3	349	4	Sequence 18869, A
21	7	1.3	389	4	Sequence 18976, A
22	7	1.3	415	4	Sequence 31590, A
23	7	1.3	434	4	Sequence 27669, A
24	7	1.3	437	2	Sequence 27235, A
25	7	1.3	475	4	Sequence 4, Appl
26	7	1.3	481	4	Sequence 4223, Ap
27	7	1.3	498	1	Sequence 30005, A
					Sequence 24, Appl

28	7	1.3	498	5	PCT-US95-05758-24	Sequence 24, Appl
29	7	1.3	507	1	US-08-363-475-19	Sequence 19, Appl
30	7	1.3	526	4	US-09-252-991A-21748	Sequence 21748, A
31	7	1.3	532	1	US-08-363-475-22	Sequence 22, Appl
32	7	1.3	544	4	US-09-198-452A-380	Sequence 380, App
33	7	1.3	565	2	US-08-700-548-2	Sequence 2, Appl
34	7	1.3	566	3	US-08-335-865-9	Sequence 9, Appl
35	7	1.3	589	4	US-09-328-352-7592	Sequence 7592, Ap
36	7	1.3	593	3	US-08-335-865-21	Sequence 21, Appl
37	7	1.3	638	4	US-09-687-538B-4	Sequence 4, Appl
38	7	1.3	833	4	US-09-252-991A-20473	Sequence 20473, A
39	7	1.3	886	4	US-09-134-001C-4496	Sequence 4496, Ap
40	7	1.3	1112	3	US-09-353-585-2	Sequence 2, Appl
41	7	1.3	1112	3	US-09-353-585-3	Sequence 3, Appl
42	7	1.3	2938	5	PCT-US94-00198-3	Sequence 3, Appl
43	6	1.1	9	2	US-08-146-028-342	Sequence 342, App
44	6	1.1	9	2	US-08-146-028-343	Sequence 343, App
45	6	1.1	9	2	US-08-146-028-344	Sequence 344, App
46	6	1.1	9	2	US-08-146-028-345	Sequence 345, App
47	6	1.1	9	2	US-08-146-028-348	Sequence 348, App
48	6	1.1	9	2	US-08-146-028-349	Sequence 349, App
49	6	1.1	9	2	US-08-146-028-350	Sequence 350, App
50	6	1.1	9	2	US-08-146-028-351	Sequence 351, App
51	6	1.1	9	3	US-08-723-425A-342	Sequence 342, App
52	6	1.1	9	3	US-08-723-425A-343	Sequence 343, App
53	6	1.1	9	3	US-08-723-425A-344	Sequence 344, App
54	6	1.1	9	3	US-08-723-425A-345	Sequence 345, App
55	6	1.1	9	3	US-08-723-425A-348	Sequence 348, App
56	6	1.1	9	3	US-08-723-425A-349	Sequence 349, App
57	6	1.1	9	3	US-08-723-425A-350	Sequence 350, App
58	6	1.1	9	3	US-08-723-425A-351	Sequence 351, App
59	6	1.1	9	3	US-09-112-206-342	Sequence 342, App
60	6	1.1	9	3	US-09-112-206-343	Sequence 343, App
61	6	1.1	9	3	US-09-112-206-344	Sequence 344, App
62	6	1.1	9	3	US-09-112-206-345	Sequence 345, App
63	6	1.1	9	3	US-09-112-206-348	Sequence 348, App
64	6	1.1	9	3	US-09-112-206-349	Sequence 349, App
65	6	1.1	9	3	US-09-112-206-350	Sequence 350, App
66	6	1.1	9	3	US-09-112-206-351	Sequence 351, App
67	6	1.1	10	3	US-08-159-339A-461	Sequence 461, App
68	6	1.1	20	2	US-08-466-975A-14	Sequence 14, Appl
69	6	1.1	20	2	US-08-466-975A-15	Sequence 15, Appl
70	6	1.1	20	2	US-08-391-671A-14	Sequence 14, Appl
71	6	1.1	20	2	US-08-391-671A-15	Sequence 15, Appl
72	6	1.1	20	3	US-08-467-902A-14	Sequence 14, Appl
73	6	1.1	20	3	US-08-467-902A-15	Sequence 15, Appl
74	6	1.1	20	3	US-09-275-265-14	Sequence 14, Appl
75	6	1.1	20	3	US-09-275-265-15	Sequence 15, Appl
76	6	1.1	20	4	US-08-850-328-10	Sequence 10, Appl
77	6	1.1	20	4	US-09-941-611-14	Sequence 14, Appl
78	6	1.1	20	4	US-09-941-611-15	Sequence 15, Appl
79	6	1.1	21	3	US-08-604-365-5	Sequence 5, Appl
80	6	1.1	22	2	US-08-146-028-55	Sequence 55, Appl
81	6	1.1	22	2	US-08-146-028-56	Sequence 56, Appl
82	6	1.1	22	2	US-08-146-028-152	Sequence 152, App
83	6	1.1	22	3	US-08-723-425A-55	Sequence 55, Appl
84	6	1.1	22	3	US-08-723-425A-56	Sequence 56, Appl
85	6	1.1	22	3	US-08-723-425A-143	Sequence 143, App
86	6	1.1	22	3	US-08-723-425A-152	Sequence 152, App
87	6	1.1	22	3	US-09-112-206-55	Sequence 55, Appl
88	6	1.1	22	3	US-09-112-206-56	Sequence 56, Appl
89	6	1.1	22	3	US-09-112-206-143	Sequence 143, App
90	6	1.1	22	3	US-09-112-206-152	Sequence 152, App
91	6	1.1	22	3	US-08-469-360A-562	Sequence 562, App
92	6	1.1	29	4	US-08-488-446-562	Sequence 562, App
93	6	1.1	29	4	US-08-467-344A-562	Sequence 562, App
94	6	1.1	29	4	US-08-146-028-57	Sequence 57, Appl
95	6	1.1	33	2	US-08-146-028-144	Sequence 144, App
96	6	1.1	33	3	US-08-723-425A-57	Sequence 57, Appl
97	6	1.1	33	3	US-08-723-425A-144	Sequence 144, App
98	6	1.1	33	3	US-08-604-365-2	Sequence 2, Appl
99	6	1.1	33	3	US-09-112-206-57	Sequence 57, Appl
100	6	1.1	33	3		

ALIGNMENTS

RESULT 1
US-09-462-606-63
; Sequence 63, Application US/09462606
; Patent No. 6432408
; GENERAL INFORMATION:
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
; FILE REFERENCE: 20264287US1
; CURRENT APPLICATION NUMBER: US/09/462,606
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 60/053069
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/14665
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-09-462-606-63

Query Match 1.5%; Score 8; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 MNNMSSAA 334
DB 1 MNNMSSAA 8
|||||

RESULT 2
US-08-363-475-1
; Sequence 1, Application US/08363475
; Patent No. 5516679
; GENERAL INFORMATION:
; APPLICANT: Chiang, Shu-Jen
; APPLICANT: Burnett Jr., William V.
; APPLICANT: Tonzi, Sean M.
; TITLE OF INVENTION: PENICILLIN V AMIDOHYDROLASE GENE FROM
; TITLE OF INVENTION: FUSARIUM OXYSPORUM
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas R. Savitsky
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,475
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitsky, Thomas R.
; REGISTRATION NUMBER: 31,661
; REFERENCE/DOCKET NUMBER: ON-0134
; TELEPHONE: (609) 252 4956
; TELEFAX: (609) 252 4526
; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Fusarium oxysporum
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 7
; OTHER INFORMATION: /note= "Amino acid 7 can be ala or
; OTHER INFORMATION: lys"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 22
; OTHER INFORMATION: /note= "Amino acid 22 can be Thr or
; OTHER INFORMATION: Val"
US-08-363-475-1

Query Match 1.3%; Score 7; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 SLKLPGT 121
DB 15 SLKLPGT 21
|||||

RESULT 3
US-09-732-210-1483
; Sequence 1483, Application US/09732210
; Patent No. 6573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 1483
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-732-210-1483

Query Match 1.3%; Score 7; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 YFRRRK 52
DB 3 YFRRRK 9
|||||

RESULT 4
US-09-732-210-1485
; Sequence 1485, Application US/09732210
; Patent No. 6573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.

; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 1485
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-732-210-1485

Query Match 1.3%; Score 7; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 YFRRRKF 52
Db 3 YFRRRKF 9

RESULT 5

US-09-732-210-1678
; Sequence 1678, Application US/09732210
; Patent No. 6573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 1678
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Sulfolobus solfataricus
US-09-732-210-1678

Query Match 1.3%; Score 7; DB 4; Length 140;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 MEPLLLA 43
Db 53 MEPLLLA 59

RESULT 6

US-09-328-352-6436
; Sequence 6436, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: CT99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6436
; LENGTH: 198

; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6436

Query Match 1.3%; Score 7; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 314 VTLLCGI 320
Db 25 VTLLCGI 31

RESULT 7

US-09-252-991A-18398
; Sequence 18398, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18398
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18398

Query Match 1.3%; Score 7; DB 4; Length 210;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 SLAAPAA 32
Db 13 SLAAPAA 19

RESULT 8

US-08-462-778-2
; Sequence 2, Application US/08462778
; Patent No. 6077517
; GENERAL INFORMATION:
; APPLICANT: Thomas, Wayne R.
; APPLICANT: Chua, Kaw-Yan
; TITLE OF INVENTION: Allergenic Protein and Peptides From
; TITLE OF INVENTION: House Dust Mite and Uses Therefor
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,778
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/031,141
; FILING DATE: 12 March 1993

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-053CP (IMI-032CP)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-227-7400
; TELEFAX: 617-227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-462-778-2

Query Match 1.3%; Score 7; DB 3; Length 215;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 400 MTLTSPF 406
DB 152 MTLTSPF 158

RESULT 9
US-08-553-336A-2
; Sequence 2, Application US/08553336A
; Patent No. 6413738
; GENERAL INFORMATION:
; APPLICANT: Wayne R. Thomas and Kaw-Yan Chua
; TITLE OF INVENTION: Allergenic Proteins and Peptides From
; TITLE OF INVENTION: House Dust Mite and Uses Therefor
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSES: LAHIVE & COCKFIELD, LLP
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,336A
; FILING DATE: 10-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/081,540
; FILING DATE: 22-JUNE-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane B. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: IMI-032CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-553-336A-2

Query Match 1.3%; Score 7; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 400 MTLTSPF 406
DB 152 MTLTSPF 158

; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-053CP (IMI-032CP)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-227-7400
; TELEFAX: 617-227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-462-778-2

Query Match 1.3%; Score 7; DB 3; Length 215;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 400 MTLTSPF 406
DB 152 MTLTSPF 158

RESULT 10
US-09-107-532A-5529
; Sequence 5529, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A. Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5529:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...244
; SEQUENCE DESCRIPTION: SEQ ID NO: 5529:
US-09-107-532A-5529

Query Match 1.3%; Score 7; DB 4; Length 244;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 TALNLFK 304
DB 191 TALNLFK 197

RESULT 11
US-08-818-112-142
; Sequence 142, Application US/08818112
; Patent No. 6290969
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
```

APPLICANT: Vedwick, Thomas S.
APPLICANT: Twardzik, Daniel R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,112
FILING DATE: 13-MAR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 142:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-818-112-142

Query Match 1.3%; Score 7; DB 3; Length 267;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 VPRPGTS 115
Db 78 VPRPGTS 84

RESULT 12
US-08-818-111-137
Sequence 137, Application US/08818111
Patent No. 6338952
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, David C.
APPLICANT: Campos-Neto, Antonia
APPLICANT: Houghton, Raymond
APPLICANT: Vedwick, Thomas S.
TITLE OF INVENTION: Twardzik, Daniel R.
NUMBER OF SEQUENCES: 148
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
TUBERCULOSIS

APPLICATION NUMBER: US/08/818,111
FILING DATE: 13-MAR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 137:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-818-111-137

Query Match 1.3%; Score 7; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 VPRPGTS 115
Db 78 VPRPGTS 84

RESULT 13
US-09-056-556-142
Sequence 142, Application US/09056556
Patent No. 6350456
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
FILING DATE: 07-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 142:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-056-556-142

Query Match 1.3%; Score 7; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 VPRPGTS 115
DB 78 VPRPGTS 84

RESULT 14
US-09-072-596-137
; Sequence 137, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 137:
; SEQUENCE CHARACTERISTICS:
; MOLECULE TYPE: peptide
; LENGTH: 267 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-072-596-137

Query Match 1.3%; Score 7; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 VPRPGTS 115
DB 78 VPRPGTS 84

RESULT 15
US-08-311-731A-173
; Sequence 173, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 173:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium leprae
; US-08-311-731A-173

Query Match 1.3%; Score 7; DB 4; Length 302;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GLTTTVI 8
DB 101 GLTTTVI 107

RESULT 16
US-09-328-352-6459
; Sequence 6459, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6459
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-6459

Query Match 1.3%; Score 7; DB 4; Length 304;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 LFKQGLD 308
DB 68 LFKQGLD 74

RESULT 17
US-09-252-991A-17986

; Sequence 17986, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17986
; LENGTH: 306
; TYPE: PR1
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17986

Query Match 1.3%; Score 7; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 97; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

Qy 470 EQARALL 476
|||||
Db 251 EQARALL 257

RESULT 18

US-09-135-639-4
; Sequence 4, Application US/09135639
; Patent No. 6004793
; GENERAL INFORMATION:
; APPLICANT: LUNNEN, KEITH D.
; APPLICANT: DALTON MICHAEL A.
; APPLICANT: WILSON, GEOFFREY G.
; APPLICANT: XU, SHUANG-YONG
; TITLE OF INVENTION: Method For Cloning And Producing The Avai Restriction
; TITLE OF INVENTION: Endonuclease In E. coli And Purification Of The
; FILE REFERENCE: Avai
; CURRENT APPLICATION NUMBER: US/09/135,639
; CURRENT FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 315
; TYPE: PR1
; ORGANISM: Anabaena variabilis
US-09-135-639-4

Query Match 1.3%; Score 7; DB 3; Length 315;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 469 VEQARAL 475
|||||
Db 39 VEQARAL 45

RESULT 19

US-09-252-991A-18869
; Sequence 18869, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18869
; LENGTH: 327
; TYPE: PR1
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18869

Query Match 1.3%; Score 7; DB 4; Length 327;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 470 EQARALL 476
|||||
Db 171 EQARALL 177

RESULT 20

US-09-252-991A-18976
; Sequence 18976, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18976
; LENGTH: 349
; TYPE: PR1
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18976

Query Match 1.3%; Score 7; DB 4; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 SLAAPAA 32
|||||
Db 302 SLAAPAA 308

RESULT 21

US-09-252-991A-31590
; Sequence 31590, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31590
; LENGTH: 389
; TYPE: PR1
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31590

Query Match 1.3%; Score 7; DB 4; Length 389;

Best Local Similarity 100.0%; Pred. No. 1.2e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 169 TARPITS 175
Db 330 TARPITS 336

RESULT 22

US-09-252-991A-27669
; Sequence 27669, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27669
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27669

Query Match 1.3%; Score 7; DB 4; Length 415;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 26 SLAAPAA 32
Db 219 SLAAPAA 225

RESULT 23

US-09-252-991A-27235
; Sequence 27235, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27235
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27235

Query Match 1.3%; Score 7; DB 4; Length 434;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 478 TASSLAP 484
Db 407 TASSLAP 413

RESULT 24

US-08-883-515-4
; Sequence 4, Application US/08883515

; Patent No. 5981836
; GENERAL INFORMATION:
; APPLICANT: Osteryoung, Katherine W
; TITLE OF INVENTION: PLANT CHLOROPLAST DIVISION GENES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,515
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 920905.90016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-883-515-4

Query Match 1.3%; Score 7; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 508 VAAQKSE 514
Db 7 VAAQKSE 13

RESULT 25

US-09-328-352-4223
; Sequence 4223, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4223
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4223

Query Match 1.3%; Score 7; DB 4; Length 475;
Best Local Similarity 100.0%; Pred. No. 1.5e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 296 PVTALNL 302
Db 264 PVTALNL 270

RESULT 26

US-09-252-991A-30005
; Sequence 30005, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30005
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30005
Query Match 1.3%; Score 7; DB 4; Length 481;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 84 VYIDEID 90
Db 237 VYIDEID 243
RESULT 27
US-08-457-274A-24
; Sequence 24, Application US/08457274A
; Patent No. 5734086
; GENERAL INFORMATION:
; APPLICANT: Scott, Jeffrey G.
; APPLICANT: Tomita, Takashi
; TITLE OF INVENTION: Cytochrome P4501pr Gene and Its Uses
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,274A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1304
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 498 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: papillio polynxnes

STRAIN:
; DEVELOPMENTAL STAGE: Adult
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
US-08-457-274A-24
Query Match 1.3%; Score 7; DB 1; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 461 VLEMRKG 467
Db 245 VLEMRKG 251
RESULT 28
PCT-US95-05758-24
; Sequence 24, Application PC/TUS9505758
; GENERAL INFORMATION:
; APPLICANT: Cornell Research Foundation, Inc.
; TITLE OF INVENTION: Cytochrome P4501pr Gene and Its
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05758
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1304
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 498 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: papillio polynxnes
STRAIN:
; DEVELOPMENTAL STAGE: Adult
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
PCT-US95-05758-24
Query Match 1.3%; Score 7; DB 5; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 461 VLEMRKG 467
Db 245 VLEMRKG 251
RESULT 29

US-08-363-475-19
; Sequence 19, Application US/08363475
; Patent No. 5516679
; GENERAL INFORMATION:
; APPLICANT: Chiang, Shu-Jen
; APPLICANT: Burnett Jr., William V.
; APPLICANT: Tonzi, Sean M.
; TITLE OF INVENTION: PENICILLIN V AMIDOHYDROLASE GENE FROM
; TITLE OF INVENTION: FUSARIUM OXYSPORUM
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas R. Savitsky
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,475
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitsky, Thomas R.
; REGISTRATION NUMBER: 31,661
; REFERENCE/DOCKET NUMBER: ON-0134
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 252 4956
; TELEFAX: (609) 252 4526
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 507 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-363-475-19

Query Match 1.3%; Score 7; DB 1; Length 507;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 SLKLPQT 121
|||||
DB 15 SLKLPQT 21

RESULT 30
US-09-252-991A-21748
; Sequence 21748, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21748
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21748

Query Match 1.3%; Score 7; DB 4; Length 526;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28 AAPAAMS 34
|||||
DB 40 AAPAAMS 46
RESULT 31
US-08-363-475-22
; Sequence 22, Application US/08363475
; Patent No. 5516679
; GENERAL INFORMATION:
; APPLICANT: Chiang, Shu-Jen
; APPLICANT: Burnett Jr., William V.
; APPLICANT: Tonzi, Sean M.
; TITLE OF INVENTION: PENICILLIN V AMIDOHYDROLASE GENE FROM
; TITLE OF INVENTION: FUSARIUM OXYSPORUM
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas R. Savitsky
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,475
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitsky, Thomas R.
; REGISTRATION NUMBER: 31,661
; REFERENCE/DOCKET NUMBER: ON-0134
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 252 4956
; TELEFAX: (609) 252 4526
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-363-475-22

Query Match 1.3%; Score 7; DB 1; Length 532;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 SLKLPQT 121
|||||
DB 40 SLKLPQT 46

RESULT 32
US-09-198-452A-380
; Sequence 380, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 380

Query Match 1.3%; Score 7; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 VEQARAL 475
|||||
DB 579 VEQARAL 585

RESULT 36

US-08-335-865J-21
; Sequence 21, Application US/08335865J
; Patent No. 6107472
; GENERAL INFORMATION:
; APPLICANT: Stacker, Steven A.; Hovens, Christopher M.,
; APPLICANT: Wilks, Andrew F.
; TITLE OF INVENTION: RECEPTOR-TYPE TYROSINE KINASE-LIKE MOLECULES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Ave
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10103

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 720 KB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: ASCII/wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/335,865J
; FILING DATE: 19-January-1995

CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU93/00210
; FILING DATE: 10-May-1993

; APPLICATION NUMBER: PL2358
; FILING DATE: 11-May-1992

; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6107472man D.
; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD-5277
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3100
; TELEFAX: (212) 752-5958

; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE: human

US-08-335-865J-21

Query Match 1.3%; Score 7; DB 3; Length 593;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 PITSSSG 178
|||||
DB 277 PITSSSG 283

RESULT 37

US-09-687-538B-4
; Sequence 4, Application US/09687538B
; Patent No. 6514739
; GENERAL INFORMATION:
; APPLICANT: Ugagawa, Hiroaki
; APPLICANT: Frandsen, Torben

; APPLICANT: Nielsen, Tom
; APPLICANT: Kauppinen, Markus
; APPLICANT: Christensen, Soeren
; TITLE OF INVENTION: Lysophospholipase
; FILE REFERENCE: 5958.210-US
; CURRENT APPLICATION NUMBER: US/09/687,538B
; CURRENT FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-09-687-538B-4

Query Match 1.3%; Score 7; DB 4; Length 638;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 LSRLNLT 205
|||||
DB 90 LSRLNLT 96

RESULT 38

US-09-252-991A-20473
; Sequence 20473, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20473
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-20473

Query Match 1.3%; Score 7; DB 4; Length 823;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 468 HVEQARA 474
|||||
DB 746 HVEQARA 752

RESULT 39

US-09-134-001C-4496
; Sequence 4496, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4496
; LENGTH: 886

;
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4496

Query Match 1.3%; Score 7; DB 4; Length 886;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 371 YRRLQW 377
Db 144 YRRLQW 150

RESULT 40

US-09-353-585-2
; Sequence 2, Application US/09353585

; Patent No. 6287865

; GENERAL INFORMATION:

; APPLICANT: Dixon, Mark S

; Jones, Jonathan DG

; Jones, David A

; TITLE OF INVENTION: Plant pathogen resistance genes and uses
; thereof

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Nixon & Vanderhye PC

; STREET: 8th Floor, 1100 No. 6287865th Glebe Road

; CITY: Arlington

; STATE: Virginia

; COUNTRY: United States of America

; ZIP: 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/353,585

; FILING DATE: 15-Jul-1999

; CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12Q

; 1/68

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/930,277

; FILING DATE: 27-OCT-1997

; APPLICATION NUMBER: PCT/GB96/00785

; FILING DATE: 01-APR-1996

; APPLICATION NUMBER: GB 9506658.5

; FILING DATE: 31-MAR-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Ms Mary J Wilson

; REGISTRATION NUMBER: 32,955

; REFERENCE/DOCKET NUMBER: 620-69

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4000

; TELEFAX: (703) 816-4100

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1112 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Tomato

; STRAIN: Cf2

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-353-585-2

Query Match

Best Local Similarity 1.3%; Score 7; DB 3; Length 1112;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 198 NLSRLNL 204
Db 336 NLSRLNL 342

RESULT 41

US-09-353-585-3

; Sequence 3, Application US/09353585

; Patent No. 6287865

; GENERAL INFORMATION:

; APPLICANT: Dixon, Mark S

; Jones, David A

; Jones, Jonathan DG

; TITLE OF INVENTION: Plant pathogen resistance genes and uses
; thereof

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Nixon & Vanderhye PC

; STREET: 8th Floor, 1100 No. 6287865th Glebe Road

; CITY: Arlington

; STATE: Virginia

; COUNTRY: United States of America

; ZIP: 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/353,585

; FILING DATE: 15-Jul-1999

; CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12Q

; 1/68

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/930,277

; FILING DATE: 27-OCT-1997

; APPLICATION NUMBER: PCT/GB96/00785

; FILING DATE: 01-APR-1996

; APPLICATION NUMBER: GB 9506658.5

; FILING DATE: 31-MAR-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Ms Mary J Wilson

; REGISTRATION NUMBER: 32,955

; REFERENCE/DOCKET NUMBER: 620-69

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4000

; TELEFAX: (703) 816-4100

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1112 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Tomato

; STRAIN: Cf2

; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-353-585-3

Query Match

Best Local Similarity 1.3%; Score 7; DB 3; Length 1112;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 198 NLSRLNL 204
Db 336 NLSRLNL 342

RESULT 42

PCT-US94-00198-3

; Sequence 3, Application PC/TUS9400198

;; GENERAL INFORMATION:
;; APPLICANT: Schering Corp.
;; TITLE OF INVENTION: RAS Associated GAP Proteins
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Schering Corp.
;; STREET: 1 Girald Farms
;; CITY: Madison
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 94304-1104
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Macintosh
;; OPERATING SYSTEM: 6.0.8
;; SOFTWARE: Microsoft Word 5.1a
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US94/00198
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/004,824
;; FILING DATE: 15-JAN-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lunn, Paul G.
;; REGISTRATION NUMBER: 32,743
;; REFERENCE/DOCKET NUMBER: DX0352 PCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (201)822-7255
;; TELEFAX: (201)822-7039
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2938 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; ORIGINAL SOURCE:
;; ORGANISM: Saccharomyces cerevisiae
PCT-US94-00198-3

Query Match 1.3%; Score 7; DB 5; Length 2938;
Best Local Similarity 100.0%; Pred. No. 8.1e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 384 LFNLLGL 390
|||
Db 1248 LFNLLGL 1254

RESULT 43
US-08-146-028-342
; Sequence 342, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 342:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid

;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-146-028-342

Query Match 1.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 LLQTAS 480
|||
Db 4 LLQTAS 9

RESULT 44
US-08-146-028-343
; Sequence 343, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 343:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-343

Query Match 1.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 LLQTAS 480
|||
Db 3 LLQTAS 8

RESULT 45
US-08-146-028-344
; Sequence 344, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 344:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-344

Query Match 1.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLOTAS 480
| | | | |
Db 2 LLOTAS 7

RESULT 46

US-08-146-028-345
; Sequence 345, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:

; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 345:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-345

Query Match 1.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLOTAS 480
| | | | |
Db 1 LLOTAS 6

RESULT 47

US-08-146-028-348
; Sequence 348, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:

; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 348:
; SEQUENCE CHARACTERISTICS:

;
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-348

Query Match 1.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLOTAS 480
| | | | |
Db 4 LLOTAS 9

RESULT 48

US-08-146-028-349
; Sequence 349, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:

; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 349:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-349

Query Match 1.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLOTAS 480
| | | | |
Db 3 LLOTAS 8

RESULT 49

US-08-146-028-350
; Sequence 350, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:

; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 350:

SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-146-028-350

Query Match 1.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 LLOTAS 480
|||
DB 2 LLOTAS 7

RESULT 50

US-08-146-028-351
Sequence 351, Application US/08146028
Patent No. 5891640

GENERAL INFORMATION:

APPLICANT: PROCESS FOR THE DETERMINATION OF PEPTIDES
TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/146,028

INFORMATION FOR SEQ ID NO: 351:

SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-146-028-351

Query Match 1.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 LLOTAS 480
|||
DB 1 LLOTAS 6

RESULT 51

US-08-723-425A-342
Sequence 342, Application US/08723425A
Patent No. 6165730

GENERAL INFORMATION:

APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
NUMBER OF SEQUENCES: 453
CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,425A
FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000

TELEFAX: 703-816-4100

INFORMATION FOR SEQ ID NO: 342:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-723-425A-342

Query Match 1.1%; Score 6; DB 3; Length 9;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 LLOTAS 480
|||
DB 4 LLOTAS 9

RESULT 52

US-08-723-425A-343
Sequence 343, Application US/08723425A
Patent No. 6165730

GENERAL INFORMATION:

APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
NUMBER OF SEQUENCES: 453
CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/723,425A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: SADOFF, B.J.

REGISTRATION NUMBER: 36,663

REFERENCE/DOCKET NUMBER: 1487-13

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4000

TELEFAX: 703-816-4100

INFORMATION FOR SEQ ID NO: 343:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear


```
; MOLECULE TYPE: peptide
US-08-723-425A-343

Query Match      1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      475 LLOTAS 480
Db      3 LLOTAS 8

RESULT 53
US-08-723-425A-344
; Sequence 344, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 344:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-723-425A-344

Query Match      1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      475 LLOTAS 480
Db      3 LLOTAS 8

RESULT 54
US-08-723-425A-345
; Sequence 345, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 344:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-723-425A-344

Query Match      1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      475 LLOTAS 480
Db      2 LLOTAS 7

RESULT 55
US-08-723-425A-348
; Sequence 348, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 344:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-723-425A-344

Query Match      1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      475 LLOTAS 480
Db      1 LLOTAS 6

RESULT 56
US-08-723-425A-348
; Sequence 348, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 345:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-723-425A-345

Query Match      1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      475 LLOTAS 480
Db      1 LLOTAS 6
```

REFERENCE/DOCKET NUMBER: 1487-13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 348:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-723-425A-348

Query Match 1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 LLQTAS 480
|||||
Db 4 LLQTAS 9

RESULT 56
US-08-723-425A-349
Sequence 349, Application US/08723425A
Patent No. 6165730
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
NUMBER OF SEQUENCES: 453
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,425A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-13
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 349:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-723-425A-349

Query Match 1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 LLQTAS 480
|||||
Db 3 LLQTAS 8

RESULT 57
US-08-723-425A-350
Sequence 350, Application US/08723425A
Patent No. 6165730
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
NUMBER OF SEQUENCES: 453
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,425A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-13
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 350:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-723-425A-350

Query Match 1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 LLQTAS 480
|||||
Db 2 LLQTAS 7

RESULT 58
US-08-723-425A-351
Sequence 351, Application US/08723425A
Patent No. 6165730
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
NUMBER OF SEQUENCES: 453
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: /09112206
APPLICATION NUMBER: US/08/723,425A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-13
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 351:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-723-425A-351

Query Match 1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLOQTAS 480
Db 1 LLOQTAS 6

RESULT 59

US-09-112-206-342
Sequence 342, Application US/09112206
Patent No. 6210903
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/112,206
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/146,028
FILING DATE:
INFORMATION FOR SEQ ID NO: 342:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-112-206-342

Query Match 1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLOQTAS 480
Db 4 LLOQTAS 9

RESULT 60
US-09-112-206-343
Sequence 343, Application US/09112206
Patent No. 6210903
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/112,206
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/146,028
FILING DATE:
INFORMATION FOR SEQ ID NO: 343:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-112-206-343

Query Match 1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLOQTAS 480
Db 3 LLOQTAS 8

RESULT 61

US-09-112-206-344
Sequence 344, Application US/09112206
Patent No. 6210903
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/112,206
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/146,028
FILING DATE:
INFORMATION FOR SEQ ID NO: 344:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-112-206-344

Query Match 1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 475 LLOTAS 480
DB 2 LLOTAS 7
|||||

RESULT 62
US-09-112-206-345
; Sequence 345, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 345:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-345

Query Match 1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 475 LLOTAS 480
DB 1 LLOTAS 6
|||||

RESULT 63
US-09-112-206-348
; Sequence 348, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:

; INFORMATION FOR SEQ ID NO: 348:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-348

Query Match 1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 475 LLOTAS 480
DB 4 LLOTAS 9
|||||

RESULT 64
US-09-112-206-349
; Sequence 349, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 349:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-349

Query Match 1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 475 LLOTAS 480
DB 3 LLOTAS 8
|||||

RESULT 65
US-09-112-206-350
; Sequence 350, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 350:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-112-206-350

Query Match 1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLOQTAS 480
Db 2 LLOQTAS 7

RESULT 66
US-09-112-206-351
; Sequence 351, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 351:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-112-206-351

Query Match 1.1%; Score 6; DB 3; Length 9;
Best Local Similarity 100.0%; Pred.No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLOQTAS 480
Db 1 LLOQTAS 6

RESULT 67
US-09-159-339A-461
; Sequence 461, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
```

```
;
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding peptides and Their
; NUMBER OF SEQUENCES: 1254
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,339A
; FILING DATE: 29-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-005030US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX:
; INFORMATION FOR SEQ ID NO: 461:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-159-339A-461

Query Match 1.1%; Score 6; DB 3; Length 10;
Best Local Similarity 100.0%; Pred.No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 161 IRTPRT 166
Db 2 IRTPRT 7

RESULT 68
US-08-466-975A-14
; Sequence 14, Application US/08466975A
; Patent No. 5910404
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; APPLICANT: POLLET, DIRK
; APPLICANT: MAERTENS, GEERT
; APPLICANT: VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
```

;;
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,975A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/391,671
; FILING DATE:
; APPLICATION NUMBER: US 07/920,286
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-466-975A-14

Query Match 1.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 LLQTAS 480
Db 12 LLQTAS 17

RESULT 69
US-08-466-975A-15
; Sequence 15, Application US/08466975A
; Patent No. 5910404
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; APPLICANT: POLLET, DIRK
; APPLICANT: MAERTENS, GEERT
; APPLICANT: VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,975A

;;
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/391,671
; FILING DATE:
; APPLICATION NUMBER: US 07/920,286
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-466-975A-15

Query Match 1.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 LLQTAS 480
Db 6 LLQTAS 11

RESULT 70
US-08-391-671A-14
; Sequence 14, Application US/08391671A
; Patent No. 5922532
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; APPLICANT: POLLET, DIRK
; APPLICANT: MAERTENS, GEERT
; APPLICANT: VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/391,671A
; FILING DATE: 21-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/920,286
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-391-671A-14

Query Match 1.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLOTAS 480
Db 12 LLOTAS 17

RESULT 71

US-08-391-671A-15
Sequence 15, Application US/08391671A
Patent No. 5922532
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
APPLICANT: FOLLET, DIRK
APPLICANT: MAERTENS, GEERT
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391.671A
FILING DATE: 21-FEB-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP91/02409
FILING DATE: 13-DEC-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-391-671A-15

Query Match 1.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 LLOTAS 480
Db 6 LLOTAS 11

RESULT 72

US-08-467-902A-14
Sequence 14, Application US/08467902A
Patent No. 6007982
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
APPLICANT: FOLLET, DIRK
APPLICANT: MAERTENS, GEERT
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,902A
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/391,671
FILING DATE:
APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP91/02409
FILING DATE: 13-DEC-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-467-902A-14

Query Match 1.1%; Score 6; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 77;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 LLQTAS 480
|||||
Db 12 LLQTAS 17

RESULT 73

US-08-467-902A-15
; Sequence 15, Application US/08467902A

; Patent No. 6007982

; GENERAL INFORMATION:

; APPLICANT: DELEYS, ROBERT J

; APPLICANT: POLLET, DIRK

; APPLICANT: MAERTENS, GEERT

; APPLICANT: VAN HEUVERSWUN, HUGO

; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF

; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHUYE P.C.

; STREET: 1100 NORTH GLEBE ROAD

; CITY: ARLINGTON

; STATE: VA

; COUNTRY: USA

; ZIP: 22201

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/467,902A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/391,671

; FILING DATE:

; APPLICATION NUMBER: US 07/920,286

; FILING DATE: 14-OCT-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/EP91/02409

; FILING DATE: 13-DEC-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP 90124241.2

; FILING DATE: 14-DEC-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: SADOFF, B.J.

; REGISTRATION NUMBER: 36,663

; REFERENCE/DOCKET NUMBER: 1487-5

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 7038164000

; TELEFAX: 7038164100

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 20 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-467-902A-15

Query Match

Best Local Similarity 1.1%; Score 6; DB 3; Length 20;

; Sequence 15, Application US/08467902A

; Patent No. 6007982

; GENERAL INFORMATION:

; APPLICANT: DELEYS, ROBERT J

; APPLICANT: POLLET, DIRK

; APPLICANT: MAERTENS, GEERT

; APPLICANT: VAN HEUVERSWUN, HUGO

; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF

; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS

RESULT 74

US-09-275-265-14

; Sequence 14, Application US/09275265

; Patent No. 6287761

; GENERAL INFORMATION:

; APPLICANT: DELEYS, ROBERT J

; APPLICANT: POLLET, DIRK

; APPLICANT: MAERTENS, GEERT

; APPLICANT: VAN HEUVERSWUN, HUGO

; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF

; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHUYE P.C.

; STREET: 1100 NORTH GLEBE ROAD

; CITY: ARLINGTON

; STATE: VA

; COUNTRY: USA

; ZIP: 22201

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/275,265

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/391,671

; FILING DATE: 21-FEB-1995

; APPLICATION NUMBER: US 07/920,286

; FILING DATE: 14-OCT-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/EP91/02409

; FILING DATE: 13-DEC-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP 90124241.2

; FILING DATE: 14-DEC-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: SADOFF, B.J.

; REGISTRATION NUMBER: 36,663

; REFERENCE/DOCKET NUMBER: 1487-5

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 7038164000

; TELEFAX: 7038164100

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 20 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-09-275-265-14

Query Match

Best Local Similarity 1.1%; Score 6; DB 3; Length 20;

; Sequence 15, Application US/09275265

; Patent No. 6287761

; GENERAL INFORMATION:

; APPLICANT: DELEYS, ROBERT J

; APPLICANT: POLLET, DIRK

; APPLICANT: MAERTENS, GEERT

; APPLICANT: VAN HEUVERSWUN, HUGO

; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF

; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS

RESULT 75

US-09-275-265-15

; Sequence 15, Application US/09275265

; Patent No. 6287761

; GENERAL INFORMATION:

; APPLICANT: DELEYS, ROBERT J

; APPLICANT: POLLET, DIRK

; APPLICANT: MAERTENS, GEERT

; APPLICANT: VAN HEUVERSWUN, HUGO

; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF

; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS

RESULT 74

US-09-275-265-14

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/275,265
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/391,671
FILING DATE: 21-FEB-1995
APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP91/02409
FILING DATE: 13-DEC-1991
APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-275-265-15

Query Match 1.1%; Score 6; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 475 LLOTAS 480
Db 6 LLOTAS 11

Search completed: February 10, 2004, 13:30:08
Job time : 23 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 13:26:36 ; Search time 40 Seconds
(without alignments)

3464.356 Million cell updates/sec

Title: US-09-821-812-5

Perfect score: 537

Sequence: 1 GGLTTTGTGRLGDRPRLS.....PDHVDTHQLIKLQHFAML 537

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SPTREMBL 23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	37.4	353	4	Q96DGH
2	201	37.4	481	4	Q8NDH9
3	201	37.4	531	4	Q8TAM2
4	88	16.4	505	11	Q9DCP7
5	88	16.4	515	11	Q8VD72
6	8	1.5	123	12	Q81870
7	8	1.5	270	17	Q8THI9
8	8	1.5	283	4	Q8WYH6
9	8	1.5	337	16	Q92MG2
10	8	1.5	510	4	Q8N3J7
11	8	1.5	772	5	Q95XK1
12	8	1.5	875	16	Q8D5P4
13	8	1.5	1517	11	Q91ZD2
14	8	1.5	5192	2	Q93TW9
15	7	1.3	59	16	Q8Y4X0
16	7	1.3	71	12	Q83097

17	1.3	75	16	Q8EAH4
18	1.3	81	16	Q98IF7
19	1.3	87	15	Q9EAZ3
20	1.3	97	12	Q9QTI1
21	1.3	110	2	Q8KLF7
22	1.3	110	2	Q93RB1
23	1.3	110	16	Q8FOX4
24	1.3	116	16	Q9RV55
25	1.3	116	17	Q9YEB1
26	1.3	122	11	Q9D363
27	1.3	135	15	Q9WU00
28	1.3	137	16	Q8FDN9
29	1.3	138	2	Q93RB2
30	1.3	144	10	Q9SBM7
31	1.3	144	16	Q8RHM0
32	1.3	147	17	Q9YEL9
33	1.3	148	12	Q91E07
34	1.3	151	16	Q9K677
35	1.3	152	6	Q9BGR6
36	1.3	157	4	Q8N9D2
37	1.3	163	13	Q9DDV1
38	1.3	166	11	Q9DAQ5
39	1.3	168	4	Q96FN9
40	1.3	168	11	Q8BHA3
41	1.3	171	5	Q9W1Q0
42	1.3	183	5	Q19841
43	1.3	203	5	Q8IT72
44	1.3	205	15	Q91VR9
45	1.3	209	5	Q9VW21
46	1.3	209	16	Q9HZK2
47	1.3	213	15	Q98XD1
48	1.3	223	16	Q9RI45
49	1.3	223	16	Q9L218
50	1.3	232	17	Q8ZXM2
51	1.3	233	16	Q8YZM8
52	1.3	234	16	Q8Z9Q3
53	1.3	238	17	Q8TJK7
54	1.3	240	5	Q8MTV3
55	1.3	245	16	Q9K8R4
56	1.3	245	16	Q8GI43
57	1.3	253	16	Q8EQ10
58	1.3	262	16	Q9RU13
59	1.3	267	16	Q8Z073
60	1.3	272	12	Q9JF43
61	1.3	272	12	Q9ICC4
62	1.3	274	10	Q80359
63	1.3	279	9	Q94M23
64	1.3	280	2	Q8VWA7
65	1.3	280	2	Q9L545
66	1.3	282	17	Q9YDR1
67	1.3	283	16	Q9YGS2
68	1.3	284	16	Q9HYX2
69	1.3	287	16	Q9CCA1
70	1.3	292	6	Q95JU6
71	1.3	295	2	Q32850
72	1.3	296	2	Q49977
73	1.3	296	16	Q8PNC9
74	1.3	303	2	Q30476
75	1.3	303	16	Q34543
76	1.3	307	10	Q9FJG2
77	1.3	309	11	Q8VFI8
78	1.3	317	16	Q9KPI7
79	1.3	318	2	Q9KHD8
80	1.3	318	15	Q71317
81	1.3	323	13	Q8QFV8
82	1.3	325	2	Q8VLK7
83	1.3	327	10	Q9AWW3
84	1.3	329	16	Q8D289
85	1.3	330	6	Q9GLY8
86	1.3	331	10	Q9SX81
87	1.3	331	16	Q9ACV3
88	1.3	337	10	Q8H9F3
89	1.3	337	16	Q8UDK1

Q8eah4 shewanela	Q98if7 rhizobium l	Q9eaz3 human immun	Q9qil1 svts2 plect	Q8klf7 rhizobium e	Q93rb1 rhodovulum	Q8fex4 corynebacte	Q9rvs5 deinococcus	Q9yeb1 aeropyrum p	Q9d363 mus musculu	Q9wu00 human immun	Q8fdn9 escherichia	Q93rb2 rhodovulum	Q9bhm7 volvox cart	Q8rhm0 fusobacteri	Q9yel9 aeropyrum p	Q9ieut cydia pomon	Q9k677 bacillus ha	Q9bgr6 macaca fasc	Q8n9d2 homo sapien	Q9ddv1 xenopus lae	Q9d4g5 mus musculu	Q9efn9 homo sapien	Q8bha3 mus musculu	Q9w1q0 drosophila	Q19841 caenorhabdi	Q8it72 neospora ca	Q9ivr9 human immun	Q9v21 drosophila	Q9hzk2 pseudomonas	Q98xd1 human immun	Q9ri45 streptomyce	Q9l218 streptomyce	Q8zxm2 pyrobaculum	Q8yzm8 raistonia e	Q8z9q3 salmonella	Q8tjk7 methanosarc	Q8mtv3 drosophila	Q9k8r4 bacillus ha	Q8gi43 bruceella su	Q8eq10 oceanobacil	Q9rul3 deinococcus	Q8z073 anabaena sp	Q9jf43 vaccinia vi	Q9icc4 vaccinia vi	Q80359 arabidopsis	Q94m23 haemophilus	Q8vwa7 streptomyce	Q9l545 streptomyce	Q9ydr1 aeropyrum p	Q9ygs2 bruceella me	Q9hyx2 pseudomonas	Q9cca1 mycobacteri	Q95ju6 macaca fasc	Q32850 mycobacteri	Q49977 mycobacteri	Q8pnc9 xanthomonas	Q30476 bacillus su	Q34543 bacillus su	Q9fjg2 arabidopsis	Q8vfi8 mus musculu	Q9kpi7 vibrio chol	Q9khd8 streptomyce	Q71317 human immun	Q8qfv8 chanos chan	Q8vlk7 escherichia	Q9aww3 oryza sativ	Q8d289 wigglewort	Q9gly8 macaca fasc	Q9sxb1 arabidopsis	Q9a6v3 caulobacter	Q8h9f3 oryza sativ	Q8udk1 agrobacteri
------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	-------------------	--------------------	---------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	---------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------

90 7 1.3 342 2 Q929N0 Q929N0 bacillus st
 91 7 1.3 342 3 O13319 O13319 metarhizium
 92 7 1.3 344 6 Q95JK4 Q95JK4 macaca fasc
 93 7 1.3 344 16 Q9PDK8 Q9PDK8 xylella fas
 94 7 1.3 348 16 Q83091 Q83091 treponema p
 95 7 1.3 357 17 Q8TUT1 Q8TUT1 methanopyru
 96 7 1.3 358 16 Q8UB09 Q8UB09 agrobacteri
 97 7 1.3 359 12 Q83036 Q83036 la france d
 98 7 1.3 361 10 Q9LNS3 Q9LNS3 arabidopsis
 99 7 1.3 363 17 Q9VOD3 Q9VOD3 pyrococcus
 100 7 1.3 365 16 Q9A2P7 Q9A2P7 caulobacter

ALIGNMENTS

RESULT 1
 Q96DG8
 ID Q96DG8 PRELIMINARY; PRT; 353 AA.
 AC Q96DG8;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to RIKEN cDNA 0610012F22 gene (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC001563; AA01563.1; -
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 5.
 DR SMART; SM00028; TPR; 5.
 FT NON_TER 1
 SQ SEQUENCE 353 AA; 40439 MW; 7B6CE1C847B8D083 CRC64;
 Query Match 37.4%; Score 201; DB 4; Length 353;
 Best Local Similarity 99.7%; Pred. No. 3.5e-200;
 Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 236 ALSTHSQYKDWKWKVQIGKCYRLGMVREAEKQFSALKQOEVDVDFLYLAKVTVSLDQ 295
 DB 52 ALSTHSQYKDWKWKVQIGKCYRLGMVREAEKQFSALKQOEVDVDFLYLAKVTVSLDQ 111
 QY 296 PVTALNLFKQGLDKFPGEVTLCCGIARIYEEMNNSSAAEYKVKLKDQNTHTVXAIACIG 355
 DB 112 PVTALNLFKQGLDKFPGEVTLCCGIARIYEEMNNSSAAEYKVKLKDQNTHTVXAIACIG 171
 QY 356 SNHFSYDQPEIALRFRRLQMGYINGQLFNNLGLCCFYAQYQDMTLTSFERALS LAENE 415
 DB 172 SNHFSYDQPEIALRFRRLQMGYINGQLFNNLGLCCFYAQYQDMTLTSFERALS LAENE 231
 QY 416 EEAADVWYNLGHVAVGIGDTNLAHQCFRLALVNNNNHAEAYNNLAVLEMRKHGHEQARAL 475
 DB 232 EEAADVWYNLGHVAVGIGDTNLAHQCFRLALVNNNNHAEAYNNLAVLEMRKHGHEQARAL 291
 QY 476 LOTASSLAPHMYEPHFNFATISDKIGDLQRSYVAAQKSEAAFPDHDVDTQHLIKQLRQHPA 535
 DB 292 LOTASSLAPHMYEPHFNFATISDKIGDLQRSYVAAQKSEAAFPDHDVDTQHLIKQLRQHPA 351
 QY 536 ML 537
 DB 352 ML 353
 RESULT 2
 Q8NDH9
 ID Q8NDH9 PRELIMINARY; PRT; 481 AA.
 AC Q8NDH9;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN DKEZP434K1118.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Ottenwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL833901; CAD38757.1; -
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 4.
 DR SMART; SM00028; TPR; 8.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 481 AA; 54477 MW; C12F59469D055E53 CRC64;
 Query Match 37.4%; Score 201; DB 4; Length 481;
 Best Local Similarity 99.7%; Pred. No. 4.5e-200;
 Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 236 ALSTHSQYKDWKWKVQIGKCYRLGMVREAEKQFSALKQOEVDVDFLYLAKVTVSLDQ 295
 DB 180 ALSTHSQYKDWKWKVQIGKCYRLGMVREAEKQFSALKQOEVDVDFLYLAKVTVSLDQ 239
 QY 296 PVTALNLFKQGLDKFPGEVTLCCGIARIYEEMNNSSAAEYKVKLKDQNTHTVXAIACIG 355
 DB 240 PVTALNLFKQGLDKFPGEVTLCCGIARIYEEMNNSSAAEYKVKLKDQNTHTVXAIACIG 299
 QY 356 SNHFSYDQPEIALRFRRLQMGYINGQLFNNLGLCCFYAQYQDMTLTSFERALS LAENE 415
 DB 300 SNHFSYDQPEIALRFRRLQMGYINGQLFNNLGLCCFYAQYQDMTLTSFERALS LAENE 359
 QY 416 EEAADVWYNLGHVAVGIGDTNLAHQCFRLALVNNNNHAEAYNNLAVLEMRKHGHEQARAL 475
 DB 360 EEAADVWYNLGHVAVGIGDTNLAHQCFRLALVNNNNHAEAYNNLAVLEMRKHGHEQARAL 419
 QY 476 LOTASSLAPHMYEPHFNFATISDKIGDLQRSYVAAQKSEAAFPDHDVDTQHLIKQLRQHPA 535
 DB 420 LOTASSLAPHMYEPHFNFATISDKIGDLQRSYVAAQKSEAAFPDHDVDTQHLIKQLRQHPA 479
 QY 536 ML 537
 DB 480 ML 481
 RESULT 3
 Q8TAM2
 ID Q8TAM2 PRELIMINARY; PRT; 531 AA.
 AC Q8TAM2;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to putative.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC026351; AA026351.1; -
 DR Genew; HGNC:20087; TTC8.
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 5.
 DR SMART; SM00028; TPR; 5.

```

SQ SEQUENCE 531 AA; 60392 MW; D8BCADCDDF8A662A CRC64;

Query Match 37.4%; Score 201; DB 4; Length 531;
Best Local Similarity 99.7%; Pred. No. 4.9e-200;
Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 236 ALSTEHSGYKDWKWKVQIGKCYRLGMYREAEKQFSKALKQEMVDTFLYLAKVYVSLDQ 295
Db 230 ALSTEHSGYKDWKWKVQIGKCYRLGMYREAEKQFSKALKQEMVDTFLYLAKVYVSLDQ 289

Qy 296 PVTALNLFKQGLDKPPEVTLGCGIARIYEMNNNSAAEYKVLKQDNTHVXAIACIG 355
Db 290 PVTALNLFKQGLDKPPEVTLGCGIARIYEMNNNSAAEYKVLKQDNTHVXAIACIG 349

Qy 356 SNHFSYDQPEALRYFLRLQMGVYNGOLFNNLGLCCFYAQOYDMLTSPERALSLENE 415
Db 350 SNHFSYDQPEALRYFLRLQMGVYNGOLFNNLGLCCFYAQOYDMLTSPERALSLENE 409

Qy 416 EEAADVYNLGHVAVGIGDTNLAHQCFLALVNNNNHAEAYNNLAVLEMRKGHVQEARAL 475
Db 410 EEAADVYNLGHVAVGIGDTNLAHQCFLALVNNNNHAEAYNNLAVLEMRKGHVQEARAL 469

Qy 476 LOTASSLAPHMYEPHFNFATISKIGDLQRSYVAAQKSEAAFPDHDVTQHLIKOLRQHPA 535
Db 470 LOTASSLAPHMYEPHFNFATISKIGDLQRSYVAAQKSEAAFPDHDVTQHLIKOLRQHPA 529

Qy 536 ML 537
Db 530 ML 531

RESULT 4
Q9DCP7 PRELIMINARY; PRT; 505 AA.
AC Q9DCP7
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 0610012F22Rik protein (Hypothetical TPR repeat containing protein).
DE 0610012F22Rik.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L. M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,
RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K. F.,
RA Suzuki H., Toyooka K., Wang K. H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;

```

```

The FANTOM Consortium,
the RIKEN Genome Exploration Research Group Phase I & II Team;
"Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK002597; BAB22218.1; -.
DR EMBL; AK081697; BAC38298.1; -.
DR MGD; MGI:1923510; 0610012F22Rik.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR; 5.
DR SMART; SM00028; TPR; 5.
KW Hypothetical protein.
SQ SEQUENCE 505 AA; 57405 MW; 9B13D8F8D9EB4F22 CRC64;

Query Match 16.4%; Score 88; DB 11; Length 505;
Best Local Similarity 100.0%; Pred. No. 2.1e-82;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 QAVRPITQAGRPTGTGFLRPSQTSGRPCTMQEQAIRTPRTAYTARPTITSSSGRFVRLGTASM 188
Db 97 QAVRPITQAGRPTGTGFLRPSQTSGRPCTMQEQAIRTPRTAYTARPTITSSSGRFVRLGTASM 156

Qy 189 LTSPDGGPFNLRLNLTYSQKPKLAKA 216
Db 157 LTSPDGGPFNLRLNLTYSQKPKLAKA 184

RESULT 5
Q8VD72 PRELIMINARY; PRT; 515 AA.
AC Q8VD72
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to RIKEN cDNA 0610012F22 gene.
GN 0610012F22Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017523; AAH17523.1; -.
DR MGD; MGI:1923510; 0610012F22Rik.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR; 5.
DR SMART; SM00028; TPR; 5.
SQ SEQUENCE 515 AA; 58439 MW; 9A8ACDF59BB641C1 CRC64;

Query Match 16.4%; Score 88; DB 11; Length 515;
Best Local Similarity 100.0%; Pred. No. 2.1e-82;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 QAVRPITQAGRPTGTGFLRPSQTSGRPCTMQEQAIRTPRTAYTARPTITSSSGRFVRLGTASM 188
Db 107 QAVRPITQAGRPTGTGFLRPSQTSGRPCTMQEQAIRTPRTAYTARPTITSSSGRFVRLGTASM 166

Qy 189 LTSPDGGPFNLRLNLTYSQKPKLAKA 216
Db 167 LTSPDGGPFNLRLNLTYSQKPKLAKA 194

RESULT 6
Q81870 PRELIMINARY; PRT; 123 AA.
AC Q81870
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE ORF 3 precursor.

```

OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1594074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Rikihisa T., Mizuno K., Win K.M.;
RT "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92335008; PubMed=1630924;
RA Aye T.T., Uchida T., Ma X.Z., Iida F., Shikata T., Zhuang H.,
RA Win K.M.;
RT "Complete nucleotide sequence of a hepatitis E virus isolated from the
RT Xinjiang epidemic (1986-1988) of China.";
RL Nucleic Acids Res. 20:3512-3512(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92271462; PubMed=1589964;
RA Fry K.E., Tam A.W., Smith M.W., Kim J.P., Luk K.-C., Young L.M.,
RA Platak M., Feldman R.A., Yun K.Y., Purdy M.A., McCaustland K.A.,
RA Bradley D.W., Reyes G.R.;
RT "Hepatitis E virus (HEV): Strain variation in the nonstructural gene
RT region encoding consensus motifs for an RNA-dependent RNA polymerase
RT and an ATP/GTP binding site.";
RL Virus Genes 6:173-185(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93348763; PubMed=8346669;
RA Bi S.-L., Purdy M.A., McCaustland K.A., Margolis H.S., Bradley D.W.;
RT "The sequence of hepatitis E virus isolated directly from a single
RT source during an outbreak in China.";
RL Virus Res. 28:233-247(1993).
DR EMBL; I08816; AAA03190.1; -;
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2; 1.
KW Signal.
FT SIGNAL 14 32 POTENTIAL.
SQ SEQUENCE 123 AA; 12588 MW; 92F59ED6B49C1755 CRC64;

Query Match 1.5%; Score 8; DB 12; Length 123;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 MNMSSAA 334
DB 1 MNMSSAA 8
|||||

RESULT 7
Q8TH19 PRELIMINARY; PRT; 270 AA.
AC Q8TH19
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Carbon-nitrogen hydrolase.
GN MA4525.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZA / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArelano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.B., Grahame D.A., Guss A.M.,
RA Hederich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE011173; AM07865.1;
DR InterPro; IPR003010; NCise/CNhydrtse.
DR Pfam; PF00795; CN hydrolase; 1.
DR PROSITE; PS0263; NITRIL_CHT_3; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 270 AA; 29944 MW; 9EDDC2A05780679F CRC64;

Query Match 1.5%; Score 8; DB 17; Length 270;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 406 ERALSALAE 413
DB 24 ERALSALAE 31
|||||

RESULT 8
Q8WYH6 PRELIMINARY; PRT; 283 AA.
ID Q8WYH6
AC Q8WYH6
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
GN PP1030.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.O., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RT "Novel human cDNA clones with function of inhibiting cancer cell
RT growth.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF193040; AAG22468.1; -;
KW Hypothetical protein.
SQ SEQUENCE 283 AA; 30596 MW; AFA62CFF78599AD6 CRC64;

Query Match 1.5%; Score 8; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 GPSLAAPA 31
DB 134 GPSLAAPA 141
|||||

```

RESULT 9
Q92MG2
ID Q92MG2 PRELIMINARY; PRT; 337 AA.
AC Q92MG2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical transmembrane protein SMC00712.
GN R02660 OR SMC00712.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Gohrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puhler A., Fumelle B., Rampsperger U.,
RA Renard C., Thebault P., Vandebol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591791; CAC47239.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 337 AA; 35599 MW; 99CFCCF915F896BF CRC64;

Query Match 1.5%; Score 8; DB 16; Length 337;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 GPSLAAPA 31
Db 201 GPSLAAPA 208
|||||

RESULT 10
Q9N3J7
ID Q9N3J7 PRELIMINARY; PRT; 510 AA.
AC Q9N3J7;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKP2P761E198.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amvgdala;
RA Koehler K., Beyer A., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834269; CAD38944.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 510 AA; 55161 MW; 29B4B5B6C350AE15 CRC64;

Query Match 1.5%; Score 8; DB 4; Length 510;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 GPSLAAPA 31
Db 264 GPSLAAPA 271
|||||

RESULT 11
Q95XK1
ID Q95XK1 PRELIMINARY; PRT; 772 AA.
AC Q95XK1;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical 87.6 kDa protein.
GN Y102A11A.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodirinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Lamar B., Fulton B., Minx P., Haakenson W., Elliott G., Gregory S.;
RT "The sequence of C. elegans cosmid Y102A11A.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC084152; AAK39314.2; -.
DR WormPep; Y102A11A.2; CE29093.
KW Hypothetical protein.
SQ SEQUENCE 772 AA; 87562 MW; 86EBE1C830C7FD1C CRC64;

Query Match 1.5%; Score 8; DB 5; Length 772;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 522 DTQHLIKQ 529
Db 725 DTQHLIKQ 732
|||||

RESULT 12
Q8D5P4
ID Q8D5P4 PRELIMINARY; PRT; 875 AA.
AC Q8D5P4;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN VV20864.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016811; AA007787.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 875 AA; 97611 MW; ABEE60479798C96 CRC64;

Query Match 1.5%; Score 8; DB 16; Length 875;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 402 LTSFERAL 409
DB 803 LTSFERAL 810

RESULT 13
Q912D2 PRELIMINARY; PRT; 1517 AA.
AC Q912D2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CCAAT displacement protein CDP.
GN CUTL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429086; PubMed=11541187;
RA Ellis T., Gambardella L., Horcher M., Tschanz S., Capol J.,
RA Bertram P., Jochum W., Barrandon Y., Busslinger M.;
RT "The transcriptional repressor CDP (Cutl1) is essential for epithelial
RT cell differentiation of the lung and the hair follicle.";
RL Gense Dev. 15:2307-2319(2001).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AY037807; RAKS9986.1; -.
DR MGD; MGI:88568; Cutl1.
DR InterPro; IPR005613; AIP3.
DR InterPro; IPR007108; Cut_homeo.
DR InterPro; IPR003350; HmoEO_CUT.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF03915; AIP3; 1.
DR Pfam; PF02376; CUT; 3.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00360; CUT_HOMEOBOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
KW SEQUENCE 1517 AA; 165852 MW; B17A8740621EBBAC CRC64;

Query Match 1.5%; Score 8; DB 11; Length 1517;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 PSLAAPAA 32
DB 1415 PSLAAPAA 1422

RESULT 14
Q93TW9 PRELIMINARY; PRT; 5192 AA.
AC Q93TW9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE MxAc.
GN MXAC.
OS Stigmatella aurantiaca.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cyctobacterineae; Cyctobacteraceae; Stigmatella.
OX NCBI_TaxID=41;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21110452; PubMed=11182319;
RA Silakowski B., Nordsiek G., Kunze B., Blocker H., Muller R.;
RT "Novel features in a combined polyketide synthase/non-ribosomal
RT peptide synthetase: the myxalamid biosynthetic gene cluster of the

```

```

RT myxobacterium Stigmatella aurantiaca Sgals.";
RL Chem. Biol. 8:59-69(2001).
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL; AF319998; AAK57187.1; -.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR004410; FabD.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR006162; Pantne_attach.
DR InterPro; IPR006163; Pp_bind.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00698; Acyl_transf; 3.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00109; ketoacyl-synt; 3.
DR Pfam; PF02801; ketoacyl-synt_C; 3.
DR Pfam; PF00550; pp-binding; 3.
DR TIGRFAMs; TIGR00128; fabD; 3.
DR PROSITE; PS00075; ACP_DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.
DR PROSITE; PS00098; THIOLEASE_1; 1.
DR Oxidoreductase; Phosphopantetheine; Transferase.
KW SEQUENCE 5192 AA; 558274 MW; 7C7C4AE414A31D4E CRC64;

Query Match 1.5%; Score 8; DB 2; Length 5192;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 SFERALS 411
DB 1335 SFERALS 1342

RESULT 15
Q8Y4X0 PRELIMINARY; PRT; 59 AA.
AC Q8Y4X0;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein lmo2309.
GN LMO2309.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Feihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001)
DR EMBL; AL591982; CAD00387.1; -.
DR Listlist; LMO02309; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 59 AA; 7168 MW; CA0871E680D08D62 CRC64;

Query Match 1.3%; Score 7; DB 16; Length 59;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 412 AENEBA 418

```



```

Db      17 AENESEA 23
|||||||
17 AENESEA 23

RESULT 16
Q83097 Q83097 PRELIMINARY; PRT; 71 AA.
AC Q83097;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE RNA-dependent RNA polymerase (fragment).
GN RDRP.
OS Leishmania RNA virus 1-13.
OC Viruses; dsRNA viruses; Totiviridae; Leishmanivirus.
OX NCBI_TaxID=39117;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95334386; PubMed=7610059;
RA Widmer G., Dooley S.;
RT "Phylogenetic analysis of Leishmania RNA virus and Leishmania suggests
ancient virus-parasite association.";
RL Nucleic Acids Res. 23:2300-2304(1995).
DR EMBL; L39069; AAC42114.1; -
KW RNA-directed RNA polymerase.
FT NON_TER 1
FT NON_TER 71
FT NON_TER 1
SQ SEQUENCE 71 AA; 8266 MW; F3A8FB7EEA8B563F CRC64;

Query Match 1.3%; Score 7; DB 12; Length 71;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 EIDVDQE 94
|||||||
Db 17 EIDVDQE 23

RESULT 17
Q8EAH4 Q8EAH4 PRELIMINARY; PRT; 75 AA.
AC Q8EAH4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ribosomal protein S18.
GN RPSR OR S03928.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OX Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015825; AAN56903.1; -
DR TIGR; SO3928; -
KW Complete proteome.
SQ SEQUENCE 75 AA; 8845 MW; A203277BC00A73B8 CRC64;

Query Match 1.3%; Score 7; DB 16; Length 75;
Best Local Similarity 100.0%; Pred. No. 83;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 YFRRKF 52
|||||||
Db 4 YFRRKF 10

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 YFRRKF 52
|||||||
Db 4 YFRRKF 10

RESULT 18
Q98IF7 Q98IF7 PRELIMINARY; PRT; 81 AA.
AC Q98IF7;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein mar2423.
GN MSR2423.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OX Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP002999; BAB49559.1; -
DR InterPro; IPR002145; HTH_CopG.
DR Pfam; PF01402; HTH_4; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 81 AA; 9035 MW; 7652A22945D0BFC7 CRC64;

Query Match 1.3%; Score 7; DB 16; Length 81;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 505 RSYVAAQ 511
|||||||
Db 29 RSYVAAQ 35

RESULT 19
Q9EAZ3 Q9EAZ3 PRELIMINARY; PRT; 87 AA.
AC Q9EAZ3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Gp120 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=A324;
RA Machuca R.A.R., Bogh M., Gerstoft J., Kvinesdal B., Pedersen C.,
RA Obel N., Nielsen H., Nielsen C.;
RT "HIV-1 subtypes in Denmark.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=A324;
RA Bogh M., Machuca R.A.R., Nielsen C.;
RT "Subtype specific problems with Roche PCR.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ289467; CAC06444.1; -

```

```
DR InterPro: IPR000777; GPI20..
DR Pfam: PF00516; GPI20; 1.
KW AID5; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 87
SQ SEQUENCE 87 AA; 9917 MW; 11415473F6861035 CRC64;

Query Match      1.3%; Score 7; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 528 KQLRQHF 534
Db 76 KQLRQHF 82

RESULT 20
Q90T11
ID Q90T11 PRELIMINARY; PRT; 97 AA.
AC Q90T11.
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF 5.
OS SVTS22 protein.
OC Viruses; ssDNA viruses; Inoviridae; Plectrovirus.
OX NCBI_TaxID=93224;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2027373; PubMed=10766306;
RA Sha Y., Melcher U., Davis R.E., Fletcher J.;
RT "Common elements of spiroplasma plectroviruses revealed by nucleotide
sequence of SVTS22."
RL Virus Genes 20:47-56(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Melcher U.K., Sha Y., Davis R.E., Fletcher J.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF133242; AAF18313.1; -.
SQ SEQUENCE 97 AA; 10656 MW; 5E1C77293FE2870B CRC64;

Query Match      1.3%; Score 7; DB 12; Length 97;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 LKLPQTN 122
Db 31 LKLPQTN 37

RESULT 21
Q8KLF7
ID Q8KLF7 PRELIMINARY; PRT; 110 AA.
AC Q8KLF7.
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein NO1E.
GN NOLE.
OS Rhizobium etli.
OG Plasmid symbiotic plasmid p42d.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=29449;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CFN42;
RA MEDLINE=91193195; PubMed=2013564;
RA Girard M.L., Flores M., Brom S., Romero D., Palacios R., Davila G.;
RT "Structural complexity of the symbiotic plasmid of Rhizobium
leguminosarum bv. phaseoli."
RL J. Bacteriol. 173:2411-2419(1991).
RN [2]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RX MEDLINE=97419521; PubMed=9274036;
RA Ramirez-Romero M.A., Bustos P., Girard L., Rodriguez O.,
RA Cevallos M.A., Davila G.;
RT "Sequence, localization and characteristics of the replicator region
of the symbiotic plasmid of Rhizobium etli.";
RL Microbiology 143:2825-2831(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RA Quintero V., Cevallos M.A., Davila G.;
RT "A site-specific recombinase and RecA are required to exert
incompatibility towards the symbiotic plasmid of Rhizobium etli.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; U80928; AAW54781.1; -.
SQ SEQUENCE 110 AA; 12009 MW; B6CDE311B94F0400 CRC64;

Query Match      1.3%; Score 7; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 AGPSLAA 29
Db 20 AGPSLAA 26

RESULT 22
Q93RB1
ID Q93RB1 PRELIMINARY; PRT; 110 AA.
AC Q93RB1.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Farnesyl diphosphate synthase (Fragment).
GN FPPS.
OS Rhodovulum strictum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodovulum.
OX NCBI_TaxID=58314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 9220;
RA Cantera J.L., Kawasaki H., Seki T.;
RT "Molecular systematic studies of phototrophic bacteria using farnesyl
diphosphate synthase."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB053177; BAB61883.1; -.
DR InterPro: IPR000092; Polyphenyl_synt.
DR Pfam: PF00348; polyphenyl_synt; 1.
FT NON_TER 1
FT NON_TER 110
SQ SEQUENCE 110 AA; 11669 MW; A715AB3622BABC27 CRC64;

Query Match      1.3%; Score 7; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 407 RALSIAE 413
Db 86 RALSIAE 92

RESULT 23
Q8FQX4
ID Q8FQX4 PRELIMINARY; PRT; 110 AA.
AC Q8FQX4.
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urease gamma subunit (EC 3.5.1.5).
GN UREA OR CE0993.
```

```

OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCB1_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayashi Y., Hino Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005217; BAC17803.1; -
KW Hydrolyase; Complete proteome.
SQ SEQUENCE 110 AA; 12415 MW; 57ED1639D3C18613 CRC64;

Query Match 1.3%; Score 7; DB 16; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 93 QEGIAEM 99
Db 80 QEGIAEM 86

RESULT 24
Q9RV55 PRELIMINARY; PRT; 116 AA.
ID Q9RV55
AC Q9RV55;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein DR0947.
GN DR0947.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCB1_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Panphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL; AB001947; AAF10529.1; -
DR TIGR; DR0947; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 116 AA; 12298 MW; 168391024730A71E CRC64;

Query Match 1.3%; Score 7; DB 16; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 468 HVEQARA 474
Db 8 HVEQARA 14

RESULT 25
Q9YEB1 PRELIMINARY; PRT; 116 AA.
ID Q9YEB1
AC Q9YEB1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

```

```

DE Hypothetical protein APE0663.
GN APE0663.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococccaeae; Aeropyrum.
OX NCB1_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K3;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1."
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA79635.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 116 AA; 12585 MW; 237AC678B52A85CF CRC64;

Query Match 1.3%; Score 7; DB 17; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 177 SGRFVRL 183
Db 20 SGRFVRL 26

RESULT 26
Q9D363 PRELIMINARY; PRT; 122 AA.
ID Q9D363
AC Q9D363;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 6530401N04Rik protein.
GN 6530401N04Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirrl L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bona M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK018311; BAB31157.1; -
DR MGD; MGI:1923485; 6530401N04Rik.
SQ SEQUENCE 122 AA; 13280 MW; 20C95DCC473558B8 CRC64;

```

Query Match 1.3%; Score 7; DB 11; Length 122;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 471 QARALLQ 477
|||||||
DB 9 QARALLQ 15

RESULT 27
Q9WIUO
ID Q9WIU0 PRELIMINARY; PRT; 135 AA.
AC Q9WIU0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Gp120 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97NOGILL;
RX MEDLINE=20092438; PubMed=10628816;
RA Jonassen T.O., Grinde B., Asjo B., Hasle G., Hungenes O.;
RT "Inter-subtype recombinant HIV-1 involving HIV-MAL-like and subtype H-
like sequence in four Norwegian cases.";
RL AIDS Res. Hum. Retroviruses 16:49-58(2000).
DR EMBL; AJ237578; CAB39753.1; -.
DR InterPro; IPR000777; Gp120.
DR Pfam; PF00516; Gp120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 135
SQ SEQUENCE 135 AA; 15100 MW; CE9BB85DB80ED7B5 CRC64;

Query Match 1.3%; Score 7; DB 15; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 KOLQHF 534
|||||||
DB 97 KOLQHF 103

RESULT 28
Q8FDN9
ID Q8FDN9 PRELIMINARY; PRT; 137 AA.
AC Q8FDN9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN C3702.
OS Escherichia coli 06.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OX Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=06:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016766; AAN82148.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 137 AA; 15621 MW; 9ACF3B157F4846FC CRC64;

Query Match 1.3%; Score 7; DB 16; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 471 QARALLQ 477
|||||||
DB 52 QARALLQ 58

RESULT 29
Q93RB2
ID Q93RB2 PRELIMINARY; PRT; 138 AA.
AC Q93RB2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Farnesyl diphosphate synthase (Fragment).
GN FPPS.
OS Rhodovulum strictum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OX Rhodobacteraceae; Rhodovulum.
OX NCBI_TaxID=58314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 9221;
RA Cantera J.L., Kawasaki H., Seki T.;
RT "Molecular systematic studies of phototrophic bacteria using farnesyl
diphosphate synthase.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB053176; BAB61882.1; -.
DR InterPro; IPR000092; Polyprenyl_synth.
DR Pfam; PF00348; polyprenyl_synth; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 14420 MW; AAA2E10C86FBA375 CRC64;

Query Match 1.3%; Score 7; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 407 RALSIAE 413
|||||||
DB 86 RALSIAE 92

RESULT 30
Q9SBM7
ID Q9SBM7 PRELIMINARY; PRT; 144 AA.
AC Q9SBM7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 15.4 kDa protein.
OS Volvox carteri f. nagariensis.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OX Volvocaceae; Volvox.
OX NCBI_TaxID=3068;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20119586; PubMed=10654090;
RA Meissner M., Stark K., Cresnar B., Kirk D.L., Schmitt R.;
RT "Volvox germline-specific genes that are putative targets of RegA
Cur. Genet. 36:363-370(1999).
RL EMBL; AF110789; AAD55571.1; -.
KW Hypothetical protein.
SQ SEQUENCE 144 AA; 15353 MW; 293F4FFA317EDF43 CRC64;

Query Match 1.3%; Score 7; DB 10; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 477 QTASSLA 483
 |||||
 Db 45 QTASSLA 51

RESULT 31

ID Q8RHMO PRELIMINARY; PRT; 144 AA.
 AC Q8RHMO;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Hypothetical protein FN1994.
 GN FN1994.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
 CC Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RX MEDLINE=21886394; PubMed=11889109;
 RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fostein M., Kyripides N., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 nucleatum strain ATCC 25586.";
 RL J. Bacteriol. 184:2005-2018(2002).
 DR EMBL; AE010501; AAL94084.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 144 AA; 17116 MW; C273FDE17F99DC2C CRC64;

Query Match 1.3%; Score 7; DB 16; Length 144;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 EMNNSS 332
 |||||
 Db 15 EMNNSS 21

RESULT 32

ID Q9VEL9 PRELIMINARY; PRT; 147 AA.
 AC Q9VEL9;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein APE0559.
 GN APE0559.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 CC Desulfurococaceae; Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kavarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,
 RA Haseyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kubota Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 DR EMBL; AP000060; BAA79527.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 147 AA; 14841 MW; 0E97973BF8C34C5 CRC64;

Query Match 1.3%; Score 7; DB 17; Length 147;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 PITSSG 178
 |||||
 Db 98 PITSSG 104

RESULT 33

ID Q91EU7 PRELIMINARY; PRT; 148 AA.
 AC Q91EU7;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE ORF108 similar to AcMNPV ORF75.
 GN ORF108.
 OS Cydia pomonella granulosis virus (CpGV) (Cydia pomonella
 granulovirus).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
 OX NCBI_TaxID=28289;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mexican 1;
 RX MEDLINE=93188168; PubMed=8445726;
 RA Crook N.E., Clem R.J., Miller L.K.;
 RA "An apoptosis-inhibiting baculovirus gene with a zinc finger-like
 motif.";
 RT J. Virol. 67:2168-2174 (1993).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mexican 1;
 RX MEDLINE=96207404; PubMed=8615018;
 RA Theilmann D.A., Chantler J.K., Stewart S., Flipsen H.T., Vlak J.M.,
 RA Crook N.E.;
 RT "Characterization of a highly conserved baculovirus structural protein
 that is specific for occlusion-derived virions.";
 RL Virology 218:148-158(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mexican 1;
 RX MEDLINE=97380577; PubMed=9237352;
 RA Kang W., Crook N.E., Winstanley D., O'Reilly D.R.;
 RT "Complete sequence and transposon mutagenesis of the BamHI J fragment
 of Cydia pomonella granulosis virus.";
 RL Virus Genes 14:131-136(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mexican 1;
 RX MEDLINE=98418511; PubMed=9747739;
 RA Kang W., Fristem M., Maeda S., Crook N.E., O'Reilly D.R.;
 RT "Identification and characterization of the Cydia pomonella
 granulovirus cathepsin and chitinase genes.";
 RL J. Gen. Virol. 79:2283-2292(1998).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mexican 1;
 RA Luque T., Finch R., Crook N., O'Reilly D.R., Winstanley D.;
 RT "The complete sequence of the Cydia pomonella granulovirus genome.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U53466; AAK70768.1; -;
 SQ SEQUENCE 148 AA; 17414 MW; 85D9ABCD83AAFC1C CRC64;

Query Match 1.3%; Score 7; DB 12; Length 148;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 200 SRLNLTK 206
 |||||
 Db 75 SRLNLTK 81

```
RESULT 34
Q9K677 PRELIMINARY; PRT; 151 AA.
ID AC Q9K677
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Hypothetical protein BH3855.
DE Hypothetical protein BH3855.
GN BH3855.
OS Bacillus halodurans.
OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
RL EMBL; AP001520; BAB07574.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 151 AA; 17118 MW; 3BC303ADC051993A CRC64;

Query Match 1.3%; Score 7; DB 16; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 DTFLYLA 287
DB 48 DTFLYLA 54

RESULT 35
Q9BGR6 PRELIMINARY; PRT; 152 AA.
ID AC Q9BGR6
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 16.5 kDa protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Frontal cortex;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB056406; BAB33064.1; -.
DR InterPro; IPR001870; B302.
DR InterPro; IPR003877; SPRY_receptor.
KW Hypothetical protein.
SQ SEQUENCE 152 AA; 16457 MW; 49788969F429B49F CRC64;

Query Match 1.3%; Score 7; DB 6; Length 152;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 PSLAAPA 31
DB 146 PSLAAPA 152

RESULT 36
Q8N9D2 PRELIMINARY; PRT; 157 AA.
ID AC Q8N9D2
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Hypothetical protein FLJ37699.
DE Hypothetical protein FLJ37699.
OS Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RA Nimmiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furiya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
KW Hypothetical protein.
SQ SEQUENCE 157 AA; 16983 MW; 754966F19A77216F CRC64;

Query Match 1.3%; Score 7; DB 4; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 ENEEEAA 419
DB 40 ENEEEAA 46

RESULT 37
Q9DDV1 PRELIMINARY; PRT; 163 AA.
ID AC Q9DDV1
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Homeobox protein Six4.1 (Fragment).
DE Six4.1.
GN Six4.1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21152917; PubMed=11231090;
RA Ghanbari H., Seo H.C., Fjose A., Brandli A.W.;
RT "Molecular cloning and embryonic expression of Xenopus Six homeobox
RT genes."
RL Mech. Dev. 101:271-277(2001).
RL EMBL; AF276994; AAG42360.1; -.
DR HSSP; P41778; 1DU6.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR007105; SIX.
DR InterPro; IPR007106; SIX_SINE_homeo.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR PROSITE; PS50553; SIX_DOMAIN; 1.
DR PROSITE; PS50554; SIX_HOMEODOMAIN; 1.
FT NON_TER 1
FT NON_TER 163
SQ SEQUENCE 163 AA; 18815 MW; 842DB2FA17A3790D CRC64;

Query Match 1.3%; Score 7; DB 13; Length 163;
```

Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 ILKARAL 80
Db 38 ILKARAL 44

RESULT 38

Q9D4Q5 ID Q9D4Q5 PRELIMINARY; PRT; 166 AA.
AC Q9D4Q5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 4930578P06Rik protein.
GN 4930578P06Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arkawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Mateuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staib F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Wombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK016300; BAB30185.1; -;
DR MGD; MGI:1923147; 4930578P06Rik.
SQ SEQUENCE 166 AA; 18051 MW; 73BA39249FC2E367 CRC64;

Query Match 1.3%; Score 7; DB 11; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 QARALLQ 477
Db 9 QARALLQ 15

RESULT 39

Q96FN9 ID Q96FN9 PRELIMINARY; PRT; 168 AA.
AC Q96FN9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similar to RIKEN cDNA 4930578P06 gene.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;

RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010618; AAH10618.1; -;
SQ SEQUENCE 168 AA; 18660 MW; F74D05F238C8D0A9 CRC64;

Query Match 1.3%; Score 7; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 QARALLQ 477
Db 9 QARALLQ 15

RESULT 40

Q8BHA3 ID Q8BHA3 PRELIMINARY; PRT; 168 AA.
AC Q8BHA3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical alpha/beta-Hydrolases structure containing protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Olfactory brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK032228; BAC27771.1; -;
DR EMBL; AK042771; BAC31360.1; -;
KW Hypothetical protein.
SQ SEQUENCE 168 AA; 18236 MW; 3CE2A5393043A5A4 CRC64;

Query Match 1.3%; Score 7; DB 11; Length 168;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 QARALLQ 477
Db 9 QARALLQ 15

RESULT 41

Q9W1Q0 ID Q9W1Q0 PRELIMINARY; PRT; 171 AA.
AC Q9W1Q0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG12227 protein (LPI0147P).
GN CG12227.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=1916006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson D.Y., Betchan P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Friese E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celnik S.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003460; AAF47006.1; -;
 DR EMBL; AY118610; AM49979.1; -;
 DR FlyBase; FBgn0034863; CGI2227;
 DR InterPro; IPR001232; Skp1.
 DR Pfam; PF01466; Skp1; 1.
 DR Pfam; PF03931; Skp1_POZ; 1.
 SQ SEQUENCE 171 AA; 19187 MW; 1CA71FD1BECDA25 CRC64;
 Query Match 1.3%; Score 7; DB 5; Length 171;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 413 ENEEEAA 419
 DB 71 ENEEEAA 77
 RESULT 42
 Q19841 ID Q19841 PRELIMINARY; PRT; 183 AA.
 AC Q19841;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 20.8 kDa protein.
 GS F27D9.4.
 ON Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode *C. elegans*: a platform for
 RT investigating biology. The *C. elegans* Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Bentley D.;
 RT "The sequence of *C. elegans* cosmid F27D9.";
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U49829; AAA93383.1; -;
 DR WormPep; F27D9.4; C502701.
 KW Hypothetical protein.
 SQ SEQUENCE 183 AA; 20782 MW; 11832E8929AFC9D CRC64;
 Query Match 1.3%; Score 7; DB 5; Length 183;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 408 ALSLAEN 414
 DB 99 ALSLAEN 105
 RESULT 43
 Q81T72 ID Q81T72 PRELIMINARY; PRT; 203 AA.
 AC Q81T72;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Microneme protein NcMIC11 precursor.
 GN MIC11.
 OS Neospora caninum.
 OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;
 OC Neospora.
 OX NCBI_TaxID=29176;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harper J.M., Zhou X.W., Carruthers V.B.;
 RT "Molecular Characterization of TgMIC11, a novel *Toxoplasma gondii*
 RT secretory protein that is cleaved into two disulfide-linked
 RT polypeptide chains.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF539703; AANI6380.1; -;
 KW Signal.
 FT SIGNAL; 1 58 POTENTIAL.
 FT CHAIN; 59 203 MICRONEME PROTEIN NCMIC11.
 SQ SEQUENCE 203 AA; 22120 MW; 61F27E9BF628938 CRC64;
 Query Match 1.3%; Score 7; DB 5; Length 203;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 270 FKSALKQ 276
 DB 72 FKSALKQ 78
 RESULT 44
 Q91VR9 ID Q91VR9 PRELIMINARY; PRT; 205 AA.
 AC Q91VR9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)


```

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE GP120 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=111676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98SE-Kol051;
RX MEDLINE=20236909; PubMed=1077151;
RA Toure-Kane C., Montavon C., Faye M., Gueye P., Sow P., Ndoye I.,
RA Gaye-Diallo A., Delaporte B., Peeters M., Mboup S.;
RT "Identification of all HIV type 1 group M subtypes in Senegal, a
RT country with low and stable seroprevalence";
RL AIDS Res. Hum. Retroviruses 16:603-609(2000).
DR EMBL; AJ272680; CAB85817.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 205
FT NON_TER 205
SQ SEQUENCE 205 AA; 22601 MW; 1A7BD8A9DEB836FB CRC64;

Query Match 1.3%; Score 7; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 528 KQLRQHF 534
Db 85 KQLRQHF 91

RESULT 45
Q9VM21
ID Q9VM21 PRELIMINARY; PRT; 209 AA.
AC Q9VM21;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CG9376 protein.
DE NCBI_TaxID=7227;
GN CG9376.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Frankkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

```

```

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL; AE003516; AAF49133.1; -.
DR FlyBase; FBGN036893; CG9376.
SQ SEQUENCE 209 AA; 23137 MW; ECA869FBB813EF8B CRC64;

Query Match 1.3%; Score 7; DB 5; Length 209;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 298 TALNLFK 304
Db 138 TALNLFK 144

RESULT 46
Q9HZK2
ID Q9HZK2 PRELIMINARY; PRT; 209 AA.
AC Q9HZK2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE Hypothetical protein PA3003.
GN PA3003.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004725; AAC06391.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 209 AA; 22474 MW; AFCAC7ACD7DF03B4 CRC64;

Query Match 1.3%; Score 7; DB 16; Length 209;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 SLAAPAA 32
Db 12 SLAAPAA 18

RESULT 47
Q98XD1
ID Q98XD1 PRELIMINARY; PRT; 213 AA.
AC Q98XD1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)

```

102 LGHVAVG 108

Db

RESULT 49

Q9L218 PRELIMINARY; PRT; 223 AA.

ID Q9L218 PRELIMINARY; PRT; 223 AA.

AC Q9L218; 223 AA.

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Putative ABC transport protein ATP-binding component.

GN SC06814 OR SC1A2.23C.

OS Streptomyces coelicolor.

OC Bacteria; Actinobacteria; Actinomycetales;

OC Streptomycetaceae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;

RN [1]

RN SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Brown S.P., Harris D.;

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

[3]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RC MEDLINE=97000351; PubMed=8843436;

RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,

RA Kinashi H., Hopwood D.A.;

RA "A set of ordered cosmids and a detailed genetic and physical map for

RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1998).

[4]

RN SEQUENCE FROM N.A.

RP STRAIN=A3(2) / M145;

RC MEDLINE=21996410; PubMed=12000953;

RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,

RA Harper D., Batenham A., Brown S., Chandra G., Chen C.W., Collins M.,

RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,

RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,

RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,

RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,

RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,

RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces

RL coelicolor A3(2).";

RL Nature 417:141-147(2002).

DR EMBL; AL939129; CAB71261.1; -;

DR InterPro; IPR003593; AAA ATPase.

DR InterPro; IPR003439; ABC transporter.

DR InterPro; IPR002453; Beta tubulin.

DR Pfam; PF00005; ABC tran; 1.

DR ProDom; PD000006; ABC transporter; 1.

DR SMART; SM00382; AAA; 1.

DR PROSITE; PS00211; ABC-TRANSPORTER; 1.

DR PROSITE; PS00228; TUBULIN B AUTOREG; 1.

DR ATP-binding; Complete proteome.

SQ SEQUENCE 223 AA; 23913 MW; D063FF42213A8694 CRC64;

Query Match 1.3%; Score 7; DB 16; Length 223;

Best Local Similarity 100.0%; Pred. No. 2.le+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 293 LDQPVT 299

Db 210 LDQPVT 216

RESULT 50

```

Q8ZXM2
ID AC Q8ZXM2 PRELIMINARY; PRT; 232 AA.
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein PAE1205.
GN PAE1205.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OK NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009809; AAL63324.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 232 AA; 25869 MW; D62F44A3849EC368 CRC64;

Query Match 1.3%; Score 7; DB 17; Length 232;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 404 SPERALS 410
Db 2 SPERALS 8

RESULT 51
Q8Y2M8
ID AC Q8Y2M8 PRELIMINARY; PRT; 233 AA.
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Putative glutathione peroxidase transmembrane protein
DE (EC 1.11.1.9).
GN RSC0307 OR RS03274.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Ralstoniaceae; Ralstonia.
OK NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646058; CAD13835.1; -.
DR InterPro; IPR000889; Glut_peroxidase.
DR Pfam; PF0255; GSPX; 1.
DR PRINTS; PR01011; GLUTPROXDAE.
DR PROSITE; PS00460; GLUTATHIONE_PEROXID 1; 1.
KW Oxidoreductase; Peroxidase; Complete proteome.
SQ SEQUENCE 233 AA; 25326 MW; 32F07B82ED7D8FE CRC64;

Query Match 1.3%; Score 7; DB 16; Length 233;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 PSLAAPA 31

```

```

Db 4 PSLAAPA 10

RESULT 52
Q8Z9Q3
ID AC Q8Z9Q3 PRELIMINARY; PRT; 234 AA.
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein STY0021.
GN STY0021.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OK NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks R., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies K.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL; AL627265; CAD01174.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 234 AA; 26524 MW; EDB40886B765E41 CRC64;

Query Match 1.3%; Score 7; DB 16; Length 234;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 177 SGRFVRL 183
Db 120 SGRFVRL 126

RESULT 53
Q8TJK7
ID AC Q8TJK7 PRELIMINARY; PRT; 238 AA.
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Glucose-1-phosphate thymidyltransferase.
GN RPB4 OR MA3777.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OK NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hederich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;

```

RT "The genome of Methanosarcina acetivorans reveals extensive metabolic

RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE011089; AAM07128.1; -.
DR InterPro; IPR005835; NTP transferase.
DR Pfam; PF00483; NTP transferase; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 238 AA; 26274 MW; C5ECA4E0B8072768 CRC64;

Query Match 1.3%; Score 7; DB 17; Length 238;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 408 ALSLAEN 414
| | | | |
DB 89 ALSLAEN 95

RESULT 54

Q8MTV3 PRELIMINARY; PRT; 240 AA.
AC Q8MTV3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE A113969p (Fragment).
GN CG9376.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dreenek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY070790; AAL48412.2; -.
DR FlyBase; FBgn0036893; CG9376.
FT NON_TER
SQ SEQUENCE 240 AA; 26740 MW; 9656D8195AB24DC9 CRC64;

Query Match 1.3%; Score 7; DB 5; Length 240;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 TALNLFK 304
| | | | |
DB 169 TALNLFK 175

RESULT 55

Q9K8R4 PRELIMINARY; PRT; 245 AA.
AC Q9K8R4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein BH2939.
GN BH2939.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hiramata C., Nakamura Y., Ogasawara N., Kuhara S.,

RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).

DR EMBL; AP001517; BAB06658.1; -.
DR InterPro; IPR002781; DUF81.
DR Pfam; PF01925; DUF81; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 245 AA; 27373 MW; 7C01109659A88691 CRC64;

Query Match 1.3%; Score 7; DB 16; Length 245;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 IGTRLGV 14
| | | | |
DB 209 IGTRLGV 215

RESULT 56

Q8G143 PRELIMINARY; PRT; 245 AA.
AC Q8G143;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN BR0880.
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122;
RA Paulsen I.T., Seehadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Unayam L., Brinkac L.M., Beanan M.J.,
RA Dougherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts."
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AE014391; AAN29808.1; -.
DR TIGR; BR0880; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 245 AA; 27613 MW; 1FE3304B831DB2B3 CRC64;

Query Match 1.3%; Score 7; DB 16; Length 245;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 VEQARAL 475
| | | | |
DB 38 VEQARAL 44

RESULT 57

Q8EQ10 PRELIMINARY; PRT; 253 AA.
AC Q8EQ10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Transposase for IS658 (divided with OB1719 and OB1720).
GN OB1719.
OS Oceanobacillus theyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.

```
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=2220767; PubMed=12233316;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus theyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments."
RL Nucleic Acids Res. 30:3927-3935(2002).
RR EMBL; AP004598; BAC13675.1; -.
RW Complete proteome.
SQ SEQUENCE 253 AA; 29749 MW; F9ADD2FCE11D728D CRC64;

Query Match 1.3%; Score 7; DB 16; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 MEQAIRT 163
Db 141 MEQAIRT 147

RESULT 59
Q9RU13 PRELIMINARY; PRT; 262 AA.
AC Q9RU13;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein DR1583.
GN DR1583.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Morfitt K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
RR EMBL; AE002002; AAF11145.1; -.
RW TIGR; DR1583; -.
SQ SEQUENCE 262 AA; 28148 MW; A3DE2D4A83C8E58E CRC64;

Query Match 1.3%; Score 7; DB 16; Length 262;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 233 IWLALST 239
Db 163 IWLALST 169

RESULT 59
Q8Z073 PRELIMINARY; PRT; 267 AA.
AC Q8Z073;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein ALR0228.
GN ALR0228.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
```

```
RN SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
RR EMBL; AP003581; BAB77752.1; -.
RW Hypothetical protein; Complete proteome.
SQ SEQUENCE 267 AA; 30918 MW; C28EA96D086E7B80 CRC64;

Query Match 1.3%; Score 7; DB 16; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 411 LAENEEE 417
Db 33 LAENEEE 39

RESULT 60
Q9JF43 PRELIMINARY; PRT; 272 AA.
AC Q9JF43;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE T88R.
OS Vaccinia virus (strain Tian Tan).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10253;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Tian Tan;
RA Jin Q., Hou Y.D., Cheng N.H., Yao E.M., Cheng S.X., Yang X.K.,
RA Jing D.Y., Yu W.H., Yuan J.S., Ma X.J.;
RT "Complete genomic sequence of vaccinia virus (Tian Tan strain)."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RR EMBL; AF095689; AAF34077.1; -.
RW InterPro; IPR000282; Cytok_receptor_2.
RR InterPro; IPR006987; Pox_IFNR.
RR Pfam; PF04903; Pox_IFNR; 1.
SQ SEQUENCE 272 AA; 31186 MW; 0A7ED9819D0DCEA6 CRC64;

Query Match 1.3%; Score 7; DB 12; Length 272;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 202 LNLTKYS 208
Db 266 LNLTKYS 272

RESULT 61
Q9ICCA PRELIMINARY; PRT; 272 AA.
AC Q9ICCA;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Interferon-gamma receptor.
GN IFNR.
OS Vaccinia virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10245;
RN [1]
RP SEQUENCE FROM N.A.
```

RC STRAIN=LIJVP;
RA Babkin I.V., Babkina I.N., Shchelkunov S.N.;
RT "Studies of variability of genes of orthopoxviruses";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ404657; CAB96934.1; -;
DR EMBL; AJ404656; CAB96933.1; -;
DR InterPro; IPR000282; Cytok receptor_2.
DR InterPro; IPR006987; Pox_IFNR.
DR Pfam; PF04903; Pox_IFNR; 1.
SQ SEQUENCE 272 AA; 31087 MW; BEFCDIC3A82DD082 CRC64;

Query Match 1.3%; Score 7; DB 12; Length 272;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 202 LNLTKYS 208
| | | | |
Db 266 LNLTKYS 272
| | | | |

RESULT 62
O80359 PRELIMINARY; PRT; 274 AA.
AC O80359;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE ATMMH-2 protein.
GN MMH-2 OR FPG2 OR F6D8.28.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Onraubo T., Matsuda O., Iba K., Terashima I., Sekiguchi M.,
RA Nakabeppu Y.;
RT "Cloning and characterization of Arabidopsis thaliana mutM orthologous
gene";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. LANDSBERG ERCTA;
RA Murphy T.M., Gao M.-J.;
RA "Two cDNAs (Accession Nos. AF099970 and AF099971) encoding Arabidopsis
RT homologs of bacterial formamidopyrimidine-DNA glycosylase genes are
RT produced by alternative processing (FGR98-204).";
RL Plant Physiol. 118:1535-1535(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA Lee J.W., Li J., Gonzalez A., Liu A., Liu K., Mukharly N., Sakano H.,
RA Vayesberg M., Chin C., Choi E., Chlou J., Altafi H., Araujo R.,
RA Brooks S., Buehler E., Chao Q., Conn L., Conway A.B., Dunn P.,
RA Hansen N., Hwang B., Huizar L., Khan S., Kim C., Palm C., Rowley D.,
RA Shinn P., Walker M., Davis R.W., Ecker J.R., Federspiel N.A.,
RA Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC F6D8 sequence";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB010690; BAA32703.1; -;
DR EMBL; AF099971; AAC97953.1; -;
DR EMBL; AC008016; AAD55613.1; -;
DR InterPro; IPR000191; Fapy_DNA_glyco.
DR Pfam; PF01149; Fapy_DNA_glyco; 1.
DR ProDom; PD003680; Fapy_DNA_glyco; 1.
DR TIGRFAMs; TIGR00577; fpy; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 274 AA; 30761 MW; E1C7829A15E0D360 CRC64;

Query Match 1.3%; Score 7; DB 10; Length 274;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 476 LQTASSL 482
| | | | |
Db 201 LQTASSL 207
| | | | |

RESULT 63
Q94WZ3 PRELIMINARY; PRT; 279 AA.
AC Q94WZ3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Orf19.
GN ORF19.
OS Haemophilus phage HP2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=157239;
RN [1]
RP SEQUENCE FROM N.A.
RA Williams B.J., Golomb M., Olson M.V., Smith A.L.;
RT "The HP2 genome of Haemophilus influenzae";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY027935; AAK37801.1; -;
SQ SEQUENCE 279 AA; 31439 MW; 2A165D5D9F380C5C CRC64;

Query Match 1.3%; Score 7; DB 9; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 407 RALSIAE 413
| | | | |
Db 103 RALSIAE 109
| | | | |

RESULT 64
Q8VWA7 PRELIMINARY; PRT; 280 AA.
AC Q8VWA7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE ACli.
GN ACli.
OS Streptomyces galilaeus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33899;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3AR-33;
RA Chung J., Fujii I., Tsukamoto N., Sankawa U., Ebizuka Y.;
RT "A flavinone-acylactinomycin biosynthesis gene cluster from Streptomyces
RT galilaeus";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008466; BAB72052.1; -;
DR InterPro; IPR005158; BAD.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF03704; BAD; 1.
DR Pfam; PF00486; trans_reg_C; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
KW DNA-binding; Sensory transduction; Transcription;
KW Transcription regulation.
SQ SEQUENCE 280 AA; 30392 MW; 0991326A0A7E175C CRC64;

Query Match 1.3%; Score 7; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 RLGMVRE 265

Db 178 RLGYRE 184
|||||||

RESULT 65

Q9L545 PRELIMINARY; PRT; 280 AA.
AC Q9L545;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE AKN1.
GN AKN1.
OS Streptomyces galliaeus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33899;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC31615;
RA Kantola J., Ylihanko K., Raty K.;
RT "A gene cluster involved in actinomycin biosynthesis from
Streptomyces galliaeus ATCC 31615.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257324; AAF70113.1; -;
DR InterPro; IPR005158; BAD.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF03704; BAD; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
SQ SEQUENCE 280 AA; 30400 MW; 01F39062011D955F CRC64;

Query Match 1.3%; Score 7; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 RLGYRE 265
|||||||

Db 178 RLGYRE 184

RESULT 66

Q9YDR1 PRELIMINARY; PRT; 282 AA.
AC Q9YDR1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein APE0856.
GN APE0856.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1.
RX MEDLINE=9310339; PubMed=10382966;
RA Kawabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
RA Hoshoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kuehida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
Crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA79836.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 282 AA; 30730 MW; 1285B52259B7E4F1 CRC64;

Query Match 1.3%; Score 7; DB 17; Length 282;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 SLAAPAA 32
|||||||

Db 22 SLAAPAA 28

RESULT 67

Q8YGS2 PRELIMINARY; PRT; 283 AA.
AC Q8YGS2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical cytosolic protein BME11086.
GN BME11086.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyripides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009548; AAL52267.1; -;
DR InterPro; IPR003768; DUF173.
DR Pfam; PF02616; DUF173; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 283 AA; 31476 MW; 55C2C1956CAB33B0 CRC64;

Query Match 1.3%; Score 7; DB 16; Length 283;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 469 VEQARAL 475
|||||||

Db 76 VEQARAL 82

RESULT 68

Q9HYX2 PRELIMINARY; PRT; 284 AA.
AC Q9HYX2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Probable transcriptional regulator.
GN PA3269.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=2043737; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- SIMILARITY: BELONGS TO THE APAC/XYLS FAMILY OF TRANSCRIPTIONAL

CC REGULATORS.
 DR EMBL; AE004749; AAG06657.1; -;
 DR InterPro; IPR000005; HTHArac.
 DR Pfam; PF00165; HTH_Arac; 2.
 DR SMART; SM00342; HTH_ARAC; 1.
 DR PROSITE; PS01124; HTH_ARAC FAMILY 2; 1.
 KW DNA-binding; Transcription regulation; Complete proteome.
 SQ SEQUENCE 284 AA; 31965 MW; E068AD8637FCEE94 CRC64;
 Query Match 1.3%; Score 7; DB 16; Length 284;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 470 EQARALL 476
 DB 229 EQARALL 235
 RESULT 69
 Q9CCAL PRELIMINARY; PRT; 287 AA.
 ID Q9CCAL
 AC Q9CCAL;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE Probable ABC-transport protein, inner membrane component.
 GN MLI088.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus."
 RL Nature 409:1007-1011(2001).
 DR EMBL; AL583920; CAC31469.1; -;
 DR LepToma; ML1088; -;
 DR InterPro; IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp; 1.
 KW Complete proteome.
 SQ SEQUENCE 287 AA; 30552 MW; 2F4115EA04123480 CRC64;
 Query Match 1.3%; Score 7; DB 16; Length 287;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GLTTTVI 8
 DB 86 GLTTTVI 92
 RESULT 70
 Q95JU6 PRELIMINARY; PRT; 292 AA.
 ID Q95JU6
 AC Q95JU6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 33.9 kDa protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Cercopithecoidea; Cercopithecoidea;

OC Cercopithecoidea; Cercopithecoidea; Cercopithecoidea;
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
 RA Terao K., Sugano S.;
 RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
 RT libraries."
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB070082; BAB63027.1; -;
 DR InterPro; IPR005045; DUF284.
 DR Pfam; PF03381; CDC50; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 292 AA; 33911 MW; BC0D893A27A6318C CRC64;
 Query Match 1.3%; Score 7; DB 6; Length 292;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 272 SALKQOE 278
 DB 16 SALKQOE 22
 RESULT 71
 Q32850 PRELIMINARY; PRT; 295 AA.
 ID Q32850
 AC Q32850;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical 29.4 kDa protein (Fragment).
 GN MB54863A.
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1765;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BCG French;
 RA Kim J.K., Choe Y.K.;
 RT "Mycobacterium bovis BCG clone E4863."
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF009829; AAB6381.1; -;
 DR InterPro; IPR002965; P-rich_extensn.
 DR PRINTS; PR01217; PRICHEXTENS.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 295 AA; 29421 MW; 04E68DF13B9A7EEC CRC64;
 Query Match 1.3%; Score 7; DB 2; Length 295;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 109 VPRPGTS 115
 DB 106 VPRPGTS 112
 RESULT 72
 Q49977 PRELIMINARY; PRT; 296 AA.
 ID Q49977
 AC Q49977;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Malg.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]


```

RP SEQUENCE FROM N.A.
RA Smith D.R.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBDJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Robison K;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBDJ databases.
DR EMBL; U15180; AAA62914.1; -.
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF00528; BPD_transp; 1.
SQ SEQUENCE 296 AA; 31610 MW; 9FE0BE849546070C CRC64;

Query Match 1.3%; Score 7; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GLTTTVI 8
Db 95 GLTTTVI 101

RESULT 73
Q8PNC9 PRELIMINARY; PRT; 296 AA.
AC Q8PNC9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Inosine-uridine preferring nucleoside hydrolase.
GN XAC1144.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Queglio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Roggi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinoza L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Teai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
DR EMBL; AE011743; AAM36016.1; -.
DR InterPro; IPR001910; I/U Nhdhase.
DR Pfam; PF01156; IU_nuc_hydro; 1.
DR ProDom; PD007736; I/U Nhdhase; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 296 AA; 31743 MW; 4FD38D7A0D2742E4 CRC64;

Query Match 1.3%; Score 7; DB 16; Length 296;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 PGEVTL 317
Db 106 PGEVTL 112

RESULT 74

```

```

O30476 PRELIMINARY; PRT; 303 AA.
AC O30476;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE YOLE.
GN YOLE.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Gim S.-Y., Jeong Y.-M., Choi S.-K., Park S.-H.;
RT "Sequence analysis of the 30 kb region (182') of the Bacillus subtilis
RT chromosome containing the cge cluster.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF006665; AAB81171.1; -.
DR InterPro; IPR000595; GMP binding.
DR InterPro; IPR004360; Gly_Eleo_diox.
DR Pfam; PF00903; Glyoxalase; 1.
DR PROSITE; PSS0042; CNMP_BINDING_3; 1.
SQ SEQUENCE 303 AA; 33823 MW; FDB061E34225F600 CRC64;

Query Match 1.3%; Score 7; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 GTSKLP 119
Db 280 GTSKLP 286

RESULT 75
O34543 PRELIMINARY; PRT; 303 AA.
AC O34543;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE YODE protein.
GN YODE.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boursier R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
RA Gim S.-Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinis S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

```

```

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler H., Yamane K., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Wambutt R., Wedler H., Lapidus A., Sorokin A., Ehrlich D.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; 299114; CAB13848.1; -.
DR EMBL; AF015775; AAB72062.1; -.
DR InterPro; IPR000595; CNMP binding.
DR InterPro; IPR004360; Gly_bleo_diox.
DR Pfam; PF00903; Glyoxalase; 1.
DR PROSITE; PS00042; CNMP_BINDING_3; 1.
KW Complete proteome.
SQ SEQUENCE 303 AA; 33851 MW; 8AB061E35B1840C2 CRC64;
Query Match 1.3%; Score 7; DB 16; Length 303;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 113 GTSLKLP 119
DB 280 GTSLKLP 286
|||||

```

Search completed: February 10, 2004, 13:28:36
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 13:26:36 ; Search time 17 Seconds
(without alignments)
1485.491 Million cell updates/sec

Title: US-09-821-812-5
Perfect score: 537
Sequence: 1 GGLTTTGTGRLGDRPRLS.....PDHVDTHQLIKQLRQHFAML 537

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	8	1.5	293	1 VTB9 AGRTU	P05358 agrobacteri
2	8	1.5	487	1 DTA STAXY	O9x2n4 staphylococ
3	8	1.5	532	1 MURD TREPA	O83873 treponema p
4	8	1.5	1395	1 CUT1 MOUSE	P53564 mus musculu
5	7	1.3	74	1 RS18 ECOLI	P02374 escherichia
6	7	1.3	74	1 RS18-HAEN	P44384 haemophilus
7	7	1.3	74	1 RS18-SALTY	P8xk81 salmonella
8	7	1.3	75	1 RS18-BUCAP	P57626 buchnera ap
9	7	1.3	75	1 RS18-BUCAP	O8k919 buchnera ap
10	7	1.3	75	1 RS18-BUCRP	P59502 buchnera ap
11	7	1.3	75	1 RS18-PASMU	P57916 pasteurella
12	7	1.3	75	1 RS18-YERPE	O8z883 yersinia pe
13	7	1.3	106	1 VCY2 HUMAN	O14599 homo sapien
14	7	1.3	110	1 NOLE RHILP	P23716 rhizobium l
15	7	1.3	133	1 Y044-BORBU	O51073 borrelia bu
16	7	1.3	137	1 RS9-SULSO	P59592 sulfolobus
17	7	1.3	137	1 RS9-SULSO	O66y33 sulfolobus
18	7	1.3	143	1 PER DROPI	Q25206 drosophila
19	7	1.3	144	1 Y244 METAC	O8tu30 methanosarc
20	7	1.3	147	1 Y872 MYCTU	Q10800 mycobacteri
21	7	1.3	174	1 YC52 PORPU	P51192 porphyra pu
22	7	1.3	215	1 ALL7 DERPT	P49273 dermatophag
23	7	1.3	223	1 GL1A ARATH	O9fma9 arabidopsis
24	7	1.3	223	1 GL1B ARATH	O9fma8 arabidopsis
25	7	1.3	229	1 LOLD VACCC	Q44613 buchnera ap
26	7	1.3	272	1 VB08 VACCC	P21004 vaccinia vi
27	7	1.3	272	1 VB08 VACCV	P24770 vaccinia vi
28	7	1.3	281	1 VPM_BHPPI	P51721 bacterioph
29	7	1.3	285	1 FMKB ECOLI	P02970 escherichia
30	7	1.3	288	1 BSN2 BACSU	O32150 bacillus su
31	7	1.3	292	1 TF RABIT	P24055 oryctolagu
32	7	1.3	315	1 T2A1 ANASP	P70803 anabaena sp
33	7	1.3	316	1 KHSE_PSEAE	P29364 pseudomonas

RESULT 1

O27848 methanobact
O91sv7 arabidopsis
P43974 haemophilus
P40937 homo sapien
P14950 columba liv
Q92040 columba liv
P53434 listeria m
Q92bq8 listeria in
Q42690 chlamydomon
Q48485 klebsiella
Q47829 erwinia her
P54913 agrobacteri
Q9y0a1 ralstonia s
O67356 aquifex aeo
Q9jtx8 neisseria m
Q9jy73 neisseria m
Q9kq24 vibrio chol
O87708 caulobacter
P33138 escherichia
Q828v1 salmonella
O8zrc0 salmonella
Q33873 yersinia en
Q8zc66 yersinia pe
Q8xyb6 ralstonia s
Q982v5 rhizobium l
Q9i2u0 pseudomonas
Q9kq87 vibrio chol
Q87r79 vibrio para
Q8dg27 vibrio vuln
Q9pe40 xylella fas
Q8k989 buchnera ap
Q9x5n1 myxococcus
Q8pn14 xanthomonas
Q8pb55 xanthomonas
P57548 buchnera ap
P33008 pseudomonas
Q04552 papilio pol
P12011 bacillus su
P04850 simian viru
P28883 simian viru
P28884 simian viru
P28885 simian viru
Q8nlh7 homo sapien
Q01887 mus musculu
Q43511 homo sapien
Q8uq90 homo sapien
Q92974 homo sapien
O13433 candida alb
Q02391 gallus gall
Q9z1e9 cricetus
Q82638 rattus norv
Q81543 mus musculu
Q92896 homo sapien
P18963 saccharomyc
Q05470 bacillus su
Q96d15 homo sapien
P80641 zea mays (m
P11736 locusta mig
P07080 bacterioph
Q96x56 sulfolobus
P08626 hordeum vul
P82977 triticum ae
Q8kuz0 vibrio chol
Q87174 vibrio para
Q8d2u3 wiggleswort
Q9hnu0 pseudomonas
Q8pml3 xanthomonas

ALIGNMENTS

VIRB9 AGRTU
ID VIRB9 AGRTU STANDARD; PRT; 293 AA.
AC P05358; P09782;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE VirB9 protein precursor.
GN VIRB9.
OS Agrobacterium tumefaciens, and
OG Agrobacterium tumefaciens (strain 15955).
OG Plasmid pTiA6, and Plasmid pTi15955.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=358, 190386;
RN [1]
RP SEQUENCE FROM N.A.
RC PLASMID=pTiA6;
RX MEDLINE=88186901; PubMed=3281947;
RA Ward J.E., Akiyoshi D.E., Regier D., Datta A., Gordon M.P.,
RA Nester E.W.;
RT "Characterization of the virB operon from an Agrobacterium
tumefaciens Ti plasmid.";
RL J. Biol. Chem. 263:5804-5814(1988).
RN [2]
RP REVISIONS.
RX MEDLINE=90170994; PubMed=2307685;
RA Ward J.E., Akiyoshi D.E., Regier D., Datta A., Gordon M.P.,
RA Nester E.W.;
RL J. Biol. Chem. 265:4768-4768(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=15955; PLASMID=pTi15955;
RX MEDLINE=88247765; PubMed=2837739;
RA Thompson D.V., Melchers L.S., Idler K.B., Shilperoort R.A.,
RA Hooykaas P.J.J.;
RT "Analysis of the complete nucleotide sequence of the Agrobacterium
tumefaciens virB operon.";
RL Nucleic Acids Res. 16:4621-4636(1988).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=15955; PLASMID=pTi15955;
RA Winans S.C., Zhu J., Oger P.M., Schrammeijer B., Hooykaas P.J.,
RA Farrand S.K.;
RT "Octopine-type Ti plasmid sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP DISULFIDE BOND, AND MUTAGENESIS OF CYS-262.
RC PLASMID=pTiA6;
RX MEDLINE=96392335; PubMed=8799123;
RA Anderson L.B., Hertz A.V., Das A.;
RT "Agrobacterium tumefaciens VirB7 and VirB9 form a disulfide-linked
protein complex.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:8889-8894(1996).
CC -!- FUNCTION: Is essential for the biogenesis of the T-pilus, which is
required for virulence and T-DNA transfer to plant cells. When is
associated with virB7, might function as a nucleation center for
recruitment of virB proteins during assembly of the T-DNA transfer
machine.
CC -!- SUBUNIT: Heterodimer of virB7 and virB9; disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Membrane-associated.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; J03216; AAA88654.1; -;
CC EMBL; X06826; CAA29979.1; -;
CC EMBL; AF242881; AAF77169.1; -;
CC PIR; S00785; B9AG55.

DR InterPro; IPR004357; IVSec_cagX.
DR Pfam; PF03524; cagX; 1.
KW Crown gall tumor; Virulence; Membrane; Signal; Plasmid.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 293 VIRB9 PROTEIN.
FT DISULFID 262 262 INTERCHAIN (WITH VIRB7).
FT MUTAGEN 262 262 C-S: NO VIRB7-VIRB9 COMPLEX FORMATION.
SQ SEQUENCE 293 AA; 32172 MW; 1BA2B48058E7DD19 CRC64;

Query Match 1.5%; Score 8; DB 1; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 293 LDQPVTAL 300
Db 173 LDQPVTAL 180
|||||||
|

RESULT 2
DLTA_STAXY STANDARD; PRT; 487 AA.
ID DLTA_STAXY
AC Q9X2N4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE D-alanine--poly(phosphoribitol)ligase subunit 1 (EC 6.1.1.13) (D-
alanine-activating enzyme) (DAE) (D-alanine-D-alanyl carrier protein
ligase) (DCL).
GN DLTA.
OS Staphylococcus xylosus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1288;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 20267 / Isolate C2A;
RX MEDLINE=99185055; PubMed=10085071;
RA Peschel A., Otto M., Jack R.W., Kalbacher H., Jung G., Gotz F.;
RT "Inactivation of the dlt operon in Staphylococcus aureus confers
sensitivity to defensins, protegrins, and other antimicrobial
peptides.";
RL J. Biol. Chem. 274:8405-8410(1999).
CC -!- FUNCTION: Involved in the biosynthesis of D-alanyl-lipoteichoic
acid (LTA). Catalyzes an ATP-dependent two-step reaction where it
forms a high energy D-alanyl AMP intermediate and transfers the
alanyl residues from AMP to Dcp (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + D-alanine + poly(ribitol phosphate) =
AMP + diphosphate + O-D-alanyl-poly(ribitol phosphate).
CC -!- PATHWAY: D-alanyl-lipoteichoic acid biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
FAMILY. DLTA SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF032440; AAD01942.1; -;
CC HSP; P14687; 1AMU.
CC HAWAP; MF_00593; -; 1.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00455; AMP_BINDING; FALSE_NEG.
KW Ligase.
SQ SEQUENCE 487 AA; 55729 MW; 283C975BC59E8BD2 CRC64;

Query Match 1.5%; Score 8; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 409 LSLAENEE 416
DB 322 LSLAENEE 329
|||||
MURD TREPA STANDARD; PRT; 532 AA.
AC O83873;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE UDP-N-acetylmuramoylalanine-D-glutamate ligase (EC 6.3.2.9) (UDP-N-
DE acetylmuramoyl-L-alanyl-D-glutamate synthetase) (D-glutamic acid
DE adding enzyme).
GN MURD OR TP0903.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Winn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Ariach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -!- FUNCTION: Cell wall formation. Catalyzes the addition of
CC glutamate to the nucleotide precursor UDP-N-acetylmuramoyl-L-
CC alanine (UMA) (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl-L-alanine +
CC Glutamate.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the murCDEF family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001259; AAC65856.1; -
DR PIR; D71267; D71267.
DR HSSP; P14900; 1EEH.
DR TIGR; TP0903; -
DR HAMAP; MF_00639; atypical; 1.
DR InterPro; IPR000713; Mur_ligase.
DR InterPro; IPR004101; Mur_ligase_C.
DR InterPro; IPR005762; MurD.
DR Pfam; PF01225; Mur_ligase; 2.
DR Pfam; PF02875; Mur_ligase_C; 1.
DR TIGRFAMs; TIGR01087; murD; 1.
DR Peptidoglycan synthetase; Cell wall; Cell division; Ligase;
KW ATP-binding; Complete proteome.
FT NP_BIND 124 130 ATP (POTENTIAL).
SQ SEQUENCE 532 AA; 57847 MW; 9FD705C2AE478ADD CRC64;

Query Match 1.5%; Score 8; DB 1; Length 532;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 470 EQARALLQ 477

```

```

DB 2 EQARALLQ 9
|||||
RESULT 4
CUT1 MOUSE STANDARD; PRT; 1395 AA.
ID CUT1_MOUSE
AC P53564; O08994; P70301;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CCAAT displacement protein (CDP) (Cut-like 1). (Homeobox protein Cux)
DE (Fragment).
GN CUTL1 OR CUX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6N;
RA Wang Z., Goldstein A., Neufeld E.J., Scheuermann R.H., Tucker P.W.;
RT "Repression of immunoglobulin heavy chain intronic enhancer
RT through nuclear matrix attachment sites: Cux/CDP homeoprotein is a
RT component of NF-muNR repressor.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 64-1395 FROM N.A. (ISOFORM 1).
RC STRAIN=A/J, and BALB/c; TISSUE=Brain;
RX MEDLINE=94244481; PubMed=7910552;
RA Valarche I., Tissier-Seta J.P., Hirsch M.R., Martinez S., Goridis C.,
RA Brunet J.F.;
RT "The mouse homeodomain protein Phox2 regulates Ncam promoter activity
RT in concert with Cux/CDP and is a putative determinant of
RT neurotransmitter phenotype.";
RL Development 119:881-896(1993).
RN [3]
RP SEQUENCE OF 642-1395 FROM N.A.
RX MEDLINE=96437626; PubMed=8840273;
RA den Heuvel G.B., Bodmer R., McConnell K.R., Nagami G.T., Igarashi P.;
RT "Expression of a cut-related homeobox gene in developing and
RT polycystic mouse kidney.";
RL Kidney Int. 50:453-461(1996).
RN [4]
RP SEQUENCE OF 936-1395 FROM N.A.
RC TISSUE=Testis;
RA Quaggin S.E., Igarashi P.;
RT "A unique variant of a homeobox gene related to Drosophila cut is
RT expressed in mouse testis.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROBABLY HAS A BROAD ROLE IN MAMMALIAN DEVELOPMENT AS A
CC REPRESSOR OF DEVELOPMENTALLY REGULATED GENE EXPRESSION. MAY ACT BY
CC PREVENTING BINDING OF POSITIVELY-ACTING CCAAT FACTORS TO
CC PROMOTERS (BY SIMILARITY). COMPONENT OF NF-MUNR REPRESSOR; BINDS
CC TO THE WARS (5' AND 3') OF THE IMMUNOGLOBULIN HEAVY CHAIN
CC ENHANCER.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P53564-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P53564-2; Sequence=VSP_002311;
CC -!- SIMILARITY: Contains 3 CUT domains.
CC -!- SIMILARITY: BELONGS TO THE CUT HOMEOBOX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```
CC -----
DR EMBL: AF004225; AAD12485.1; -
DR EMBL: X75013; CA52922.1; -
DR EMBL: U46683; AAC52775.1; -
DR EMBL: U46684; AAB41146.1; -
DR PIR: I48314; I48314.
DR HSP: P10037; 1AU7.
DR MGD: MGI:88568; Cut11.
DR GO: GO:0005634; C:nucleus; IDA.
DR GO: GO:0016564; F:transcriptional repressor activity; IDA.
DR GO: GO:0042491; P:hair cell differentiation; IMP.
DR GO: GO:0030324; P:lung development; IMP.
DR GO: GO:0000122; P:negative regulation of transcription from P...; IDA.
DR InterPro: IPR007108; Cut_hmoeo.
DR InterPro: IPR003350; Hmoeo_CUT.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF02376; CUT; 3.
DR Pfam: PF00046; homeobox; 1.
DR ProDom: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS00071; HOMEBOX_2; 1.
DR Transcription regulation; Homeobox; DNA-binding;
KW Developmental protein; Nuclear protein; Repeat; Repressor;
KW Coiled coil; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 243 COILED COIL (POTENTIAL).
FT DNA_BIND 420 507 CUT 1.
FT DOMAIN 547 603 COILED COIL (POTENTIAL).
FT DNA_BIND 809 896 CUT 2.
FT DNA_BIND 992 1079 CUT 3.
FT DNA_BIND 1119 1178 HOMEBOX.
FT VARSPIC 287 388 Missing (in isoform 2).
FT CONFLICT 1360 1360 /FTIG-VSP 002311.
FT CONFLICT 1365 1365 G -> A (IN REF. 2).
FT CONFLICT 1395 1395 P -> L (IN REF. 1).
SQ SEQUENCE 1395 AA; 151802 MW; D062CC227D7A163E CRC64;

Query Match 1.5%; Score 8; DB 1; Length 1395;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 PSLAAPAA 32
Db 1293 PSLAAPAA 1300

RESULT 5
RS18_ECOLI
ID RS18_ECOLI STANDARD; PRT; 74 AA.
AC P02374;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 30S ribosomal protein S18.
GN RPSR OR B4202 OR C5292 OR Z5811 OR ECS5178 OR STY4749 OR T4444 OR
SF4355.
OS Escherichia coli,
OS Escherichia coli O6,
OS Escherichia coli O157:H7,
OS Salmonella typhi, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 83334, 601, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=86310297; PubMed=3528756;
RA Schnier J., Kitakawa M., Isono K.;
RT "The nucleotide sequence of an Escherichia coli chromosomal region
RT containing the genes for ribosomal proteins S6, S18, L9 and an open
RT reading frame.";
```

```
RL Mol. Gen. Genet. 204:126-132 (1986).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119 (1995).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai J., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
RN [6]
RP SEQUENCE.
RC SPECIES=E.coli; STRAIN=K;
RX MEDLINE=76210737; PubMed=776663;
RA Yaguchi M.;
RT "Primary structure of protein S18 from the small Escherichia coli
RT ribosomal subunit.";
RL FEBS Lett. 59:217-220 (1975).
RN [7]
RP SEQUENCE OF 30-37, AND CROSS-LINKING TO RRNA.
RC SPECIES=E.coli; STRAIN=MRE-600;
RX MEDLINE=96003638; PubMed=7556101;
RA Uriaub H., Kruft V., Bischof O., Mueller E.-C., Wittmann-Liebold B.;
RT "Protein-rRNA binding features and their structural and functional
RT implications in ribosomes as determined by cross-linking studies.";
RL EMBO J. 14:4578-4588 (1995).
RN [8]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogan A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
```

```

RT enterica serovar Typhi CT18." ;
RL Nature 413:848-852(2001).
RN [9]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18." ;
RL J. Bacteriol. 185:2330-2337(2003).
RN [10]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157." ;
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [11]
RP MASS SPECTROMETRY.
RC SPECIES=E.coli; STRAIN=K12 / ATCC 25404;
RX MEDLINE=99196679; PubMed=10094780;
RA Arnold R.J., Reilly J.P.;
RT "Observation of Escherichia coli ribosomal proteins and their
RT posttranslational modifications by mass spectrometry." ;
RL Anal. Biochem. 269:105-112(1999).
RN [12]
RP 3D-STRUCTURE MODELING.
RC SPECIES=E.coli;
RX MEDLINE=22239879; PubMed=12244297;
RA Tung C.-S., Joseph S., Sanbonmatsu K.Y.;
RT "All-atom homology model of the Escherichia coli 30S ribosomal
RT subunit." ;
RL Nat. Struct. Biol. 9:750-755(2002).
CC -!- FUNCTION: This protein has been implicated in aminoacyl-transfer
CC RNA binding. It appears to be situated at the decoding site of
CC messenger RNA.
CC -!- MASS SPECTROMETRY: MW=8897.0; METHOD=MALDI.
CC -!- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04022; CAA27654.1; -
DR EMBL; U14003; AAA97098.1; -
DR EMBL; AE000491; AAC77159.1; -
DR EMBL; AE016771; AAN83713.1; -
DR EMBL; AE005652; AAG59398.1; -
DR EMBL; AP002568; BAB38601.1; -
DR EMBL; AL627283; CAD06870.1; -
DR EMBL; AE016849; AA071891.1; -
DR EMBL; AE015442; AAN45772.1; -
DR PIR; B86117; B86117.
DR PIR; B91276; B91276.
DR PIR; S56427; R3EC18.
DR PDB; 1M5G; 09-JUL-02.
DR EcGene; EG10917; rpeR.
DR HAMAP; MF 00270; -, 1
DR InterPro; IPR001648; Ribosomal S18.
DR Pfam; PF01084; Ribosomal S18; 1.
DR PRINTS; PR00974; RIBOSOMALS18.
DR ProDom; PD002239; Ribosomal S18; 1.
DR TIGRfams; TIGR00165; S18; 1.

```

```

DR PROSITE; PS00057; RIBOSOMAL_S18; 1.
KW Ribosomal protein; RNA-binding; Acetylation; 3D-structure;
KW Complete proteome.
FT INIT_MET 0 0 ACETYLATION
FT MOD_RES 1 1 E -> Q (IN REF. 6).
FT CONFLICT 15 15 K -> E (IN REF. 4).
FT CONFLICT 23 23
SQ SEQUENCE 74 AA; 8855 MW; AB2EDDEEA9441581 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 74;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 YFRRKRF 52
Db 3 YFRRKRF 9
|||||
RESULT 6
RS18 HAEIN STANDARD; PRT; 74 AA.
ID RS18 HAEIN STANDARD; PRT; 74 AA.
AC P44384;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S18.
GN RPSR OR RPS18 OR HI0545.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervatage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Corton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd." ;
RL Science 269:496-512(1995).
CC -!- FUNCTION: This protein has been implicated in aminoacyl-transfer
CC RNA binding. It appears to be situated at the decoding site of
CC messenger RNA (By similarity).
CC -!- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32736; AAC22203.1; -
DR PIR; E64076; E64076.
DR TIGR; HI0545; -, 1
DR HAMAP; MF 00270; -, 1
DR InterPro; IPR001648; Ribosomal S18.
DR Pfam; PF01084; Ribosomal S18; 1.
DR PRINTS; PR00974; RIBOSOMALS18.
DR ProDom; PD002239; Ribosomal S18; 1.
DR TIGRfams; TIGR00165; S18; 1.
DR PROSITE; PS00057; RIBOSOMAL_S18; 1.
DR Ribosomal protein; Acetylation; RNA-binding; Complete proteome.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).

```

```

SQ SEQUENCE 74 AA; 8811 MW; AC39A87BC54D1C81 CRC64;
Query Match 1.3%; Score 7; DB 1; Length 74;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

OY 46 YFRRRK 52
DB 3 YFRRRK 9

RESULT 7
RS18_BUCAL
ID RS18_SALTY STANDARD; PRT; 74 AA.
AC Q8ZK81,
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S18.
GN RPSR OR STW4393.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
CC -!- FUNCTION: This protein has been implicated in aminoacyl-transfer
CC RNA binding. It appears to be situated at the decoding site of
CC messenger RNA (By similarity).
CC -!- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE008906; AAL23213.1; -.
CC StyGene; SG7777; rpsR.
CC HAMAP; MF 00270; -.
CC InterPro; IPR001648; Ribosomal_S18.
CC Pfam; PF01084; Ribosomal_S18.
CC PRINTS; PR00974; RIBOSOMALS18.
CC ProDom; PD002239; Ribosomal_S18; 1.
CC TIGRFAMs; TIGR00165; S18; 1.
CC PROSITE; PS00057; RIBOSOMAL_S18; 1.
CC RIBOSOMAL protein; RNA-binding; Acetylation; Complete proteome.
FT INIT MET 1 0 ACETYLATION (BY SIMILARITY).
FT MOD_RES 1 1
SQ SEQUENCE 74 AA; 8940 MW; AB2EDDF173DA1581 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 74;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 YFRRRK 52
DB 3 YFRRRK 9

RESULT 8

```

```

RS18_BUCAL
ID RS18_BUCAL STANDARD; PRT; 75 AA.
AC P57626,
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S18.
GN RPSR OR BU563.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS."
RL Nature 407:81-86(2000).
CC -!- FUNCTION: This protein has been implicated in aminoacyl-transfer
CC RNA binding. It appears to be situated at the decoding site of
CC messenger RNA (By similarity).
CC -!- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP001119; BAB13253.1; -.
CC HAMAP; MF 00270; -.
CC InterPro; IPR001648; Ribosomal_S18.
CC Pfam; PF01084; Ribosomal_S18; 1.
CC PRINTS; PR00974; RIBOSOMALS18.
CC ProDom; PD002239; Ribosomal_S18; 1.
CC TIGRFAMs; TIGR00165; S18; 1.
CC PROSITE; PS00057; RIBOSOMAL_S18; 1.
CC RIBOSOMAL protein; RNA-binding; Complete proteome.
SQ SEQUENCE 75 AA; 8984 MW; 445E82A691FB73BA CRC64;

Query Match 1.3%; Score 7; DB 1; Length 75;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 YFRRRK 52
DB 4 YFRRRK 10

RESULT 9
RS18_BUCAL
ID RS18_BUCAL STANDARD; PRT; 75 AA.
AC Q8K919;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S18.
GN RPSR OR BUSG543.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22084549; PubMed=12089438;
RA Tamas I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;
RT "50 million years of genomic stasis in endosymbiotic bacteria."

```


Science 296:2376-2379(2002).
-!- FUNCTION: This protein has been implicated in aminoacyl-transfer RNA binding. It appears to be situated at the decoding site of messenger RNA (by similarity).

-!- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AE014128; AAM68082.1; --
HMAP; MF_00270; --; 1.
InterPro; IPR001648; Ribosomal S18.
Pfam; PF01084; Ribosomal S18; 1.
PRINTS; PR00974; RIBOSOMALS18.
ProDom; PD002239; Ribosomal S18; 1.
TIGRFAMs; TIGR00165; S18; 1.
PROSITE; PS00057; RIBOSOMAL_S18; 1.
Ribosomal protein; RNA-binding; Complete proteome.
SEQUENCE 75 AA; 8994 MW; 9578DBB68E46C8B CRC64;

Query Match 1.3%; Score 7; DB 1; Length 75;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 YFRRRK 52
| | | | |
Db 4 YFRRRK 10

RESULT 10
RS18_BUCBP ID RS18_BUCBP STANDARD; PRT; 75 AA.
AC P59502;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 30S ribosomal protein S18.
GN RPSR OR BAP509.
OS Buchnera aphidicola (subsp. Baizongia pistaciae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=1335842;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426901; PubMed=12522265;
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F., Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J., Tanames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
RT "Reductive genome evolution in Buchnera aphidicola."
RL proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC -!- FUNCTION: This protein has been implicated in aminoacyl-transfer RNA binding. It appears to be situated at the decoding site of messenger RNA (by similarity).

-!- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AE014017; AAO27212.1; --
HMAP; MF_00270; --; 1.
PROSITE; PS00057; RIBOSOMAL_S18; 1.
Ribosomal protein; RNA-binding; Complete proteome.
SEQUENCE 75 AA; 9105 MW; 0C618E14BD27F4B1 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 75;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 YFRRRK 52
| | | | |
Db 4 YFRRRK 10

RESULT 11
RS18_PASMU ID RS18_PASMU STANDARD; PRT; 75 AA.
AC P57916;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S18.
GN RPSR OR RPS18 OR PM1178.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Pm70;
MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- FUNCTION: This protein has been implicated in aminoacyl-transfer RNA binding. It appears to be situated at the decoding site of messenger RNA (by similarity).

-!- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AE006158; AAK03262.1; --
HMAP; MF_00270; --; 1.
InterPro; IPR001648; Ribosomal S18.
Pfam; PF01084; Ribosomal S18; 1.
PRINTS; PR00974; RIBOSOMALS18.
ProDom; PD002239; Ribosomal S18; 1.
TIGRFAMs; TIGR00165; S18; 1.
PROSITE; PS00057; RIBOSOMAL_S18; 1.
Ribosomal protein; RNA-binding; Complete proteome.
SEQUENCE 75 AA; 8942 MW; AC284AAB36C27AB8 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 75;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 YFRRRK 52
| | | | |
Db 4 YFRRRK 10

RESULT 12
RS18_YERPE ID RS18_YERPE STANDARD; PRT; 75 AA.
AC O82B83;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S18.
GN RPSR OR YPO3537 OR Y0646.
OS Yersinia pestis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OX Enterobacteriaceae; Yersinia.
RN NCBI_TaxID=632;
RP [1]
RC SEQUENCE FROM N.A.
RX STRAIN=CO-92 / Biovar Orientalis;
RM MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
Fellwell T., Hamlin N., Hollroyd S., Jagels K., Karlyshev A.V.,
Leather S., Moulé S., Oyston P.C.F., Quail M., Rutherford K.,
Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RL "Genome sequence of Yersinia pestis, the causative agent of plague."
RM Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
Perry R.D.;
RM "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
RC -!- FUNCTION: This protein has been implicated in aminoacyl-transfer
RNA binding. It appears to be situated at the decoding site of
messenger RNA (By similarity).
CC -!- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC
DR EMBL; AJ414157; CAC92766.1; -.
DR EMBL; AE013667; AAM84234.1; -.
DR PIR; AB0430; AB0430.
DR HAMAP; MF_00270; -; 1.
DR InterPro; IPR001648; Ribosomal_S18.
DR Pfam; PF01084; Ribosomal_S18; 1.
DR PRINTS; PR00974; RIBOSOMALS18.
DR ProDom; PD002239; Ribosomal_S18; 1.
DR TIGRFAMs; TIGR00165; S18; 1.
DR PROSITE; PS00057; RIBOSOMAL_S18; 1.
KW Ribosomal protein; RNA-binding; Complete proteome.
SQ SEQUENCE 75 AA; 8989 MW; AC2A346F971B73B8 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 75;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 YFRRRK 52
DB 4 YFRRRK 10

RESULT 13
VCY2 HUMAN
ID VCY2 HUMAN STANDARD; PRT; 106 AA.
AC O14599;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Testis-specific basic protein Y 2 (Variably charged protein Y 2).
GN VCY2 OR BPY2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98022381; PubMed=9381176;
RA Lahn B.T., Page D.C.;
RM "Functional coherence of the human Y chromosome."
RL Science 278:675-680(1997).
CC -!- TISSUE SPECIFICITY: Expressed exclusively in testis.
CC -!- SIMILARITY: BELONGS TO THE VCX/VCY FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC
DR EMBL; AF000980; AAC51828.1; -.
DR Genbank; HGNC:13508; VCX2.
DR MIM; 400013; -.
DR GO; GO:0007338; P:fertilization (sensu Animalia); TAS.
SQ SEQUENCE 106 AA; 12035 MW; 2573EC02A2DEA788 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 311 PGEVTL 317
DB 77 PGEVTL 83

RESULT 14
NOLE RHLP
ID NOLE RHLP STANDARD; PRT; 110 AA.
AC P23716;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Modulation protein nole precursor.
GN NOLE.
OS Rhizobium leguminosarum (biovar phaseoli).
OC Plasmid sym.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=385;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8002;
RX MEDLINE=91014692; PubMed=2215216;
RA Davis E.O., Johnston A.W.B.;
RM "Analysis of three nodD genes in Rhizobium leguminosarum biovar
phaseoli; nodD1 is preceded by noIE, a gene whose product is secreted
from the cytoplasm."
RL Mol. Microbiol. 4:921-932(1990).
CC -!- SUBCELLULAR LOCATION: Periplasmic (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC
DR EMBL; X54214; CAA38125.1; -.
DR PIR; S11786; S11786.
KW Plasmid; Modulation; Periplasmic; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 110 MODULATION PROTEIN NOLE.
SQ SEQUENCE 110 AA; 12050 MW; EF2FB42A7BDBC904 CRC64;

```

Query Match          1.3%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 23 AGPSLAA 29
Db 20 AGPSLAA 26

RESULT 15
Y044 BORBU          STANDARD; PRT; 133 AA.
ID Y044 BORBU
AC O51073;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 401, Last annotation update)
DE Hypothetical protein BB0044.
GN BB0044.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 35210 / B31;
MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Caajens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson S.,
RA Peterson J., Kerlavage A.R., Quackenbush J.D., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Wathley L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001118; AAC66442.1; -.
DR PIR; D70105; D70105.
DR TIGR; BB0044; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 11 31 POTENTIAL.
SQ SEQUENCE 133 AA; 16052 MW; A57686EA30F1959B CRC64;

Query Match          1.3%; Score 7; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 60 TOMLEKS 66
Db 31 TOMLEKS 37

RESULT 16
RS9_SULSO          STANDARD; PRT; 137 AA.
ID RS9_SULSO
AC P95592;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S9P.
GN RPS9P OR RPS9AB OR RPS9 OR SSO0068 OR C05003.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

Query Match          1.3%; Score 7; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 37 MEPLLLA 43
Db 50 MEPLLLA 56

RESULT 17
RS9_SULSO          STANDARD; PRT; 137 AA.
ID RS9_SULSO
AC Q96YW3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S9P.
GN RPS9P OR ST2064.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JCM 10545 / 7;
MEDLINE=21456156; PubMed=11572479;
RA Kwarababayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,

```

RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Ogunchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.,
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.",
RL DNA Res. 8:123-140(2001).
CC -!- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP000988; BAB67163.1; --
DR HAMAP; MF_00532; --; 1.
DR InterPro; IPR000754; Ribosomal_S9.
DR Pfam; PF00380; Ribosomal_S9; 1.
DR ProDom; PD001627; Ribosomal_S9; 1.
DR PROSITE; PS00360; RIBOSOMAL_S9; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 137 AA; 15723 MW; 356170BFA052C06 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 MEPLLLA 43
| | | | |
DB 50 MEPLLLA 56

RESULT 18
PER_DROP1 STANDARD; PRT; 143 AA.
AC Q25206;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Period circadian protein (fragment).
GN PER.
OS Drosophila picticornis (Fruit fly) (Idiomya picticornis).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7235;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95115533; PubMed=7815924;
RA Nielsen J., Peixoto A.A., Piccin A., Costa R., Kyriacou C.P.,
RA Chalmers D.;
RT "Big flies, small repeats: the 'Thr-Gly' region of the period gene in
RT Diptera.";
RL Mol. Biol. Evol. 11:839-853(1994).
CC -!- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE
CC PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS; AN INCREASE IN
CC PER DOSAGE LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE
CC LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN
CC RHYTHMICITY OF LOCOMOTOR ACTIVITY, ECLOSION BEHAVIOR, AND FOR THE
CC RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN
CC THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE OF THE PER-
CC RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER
CC COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES
CC ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERODIMER
CC COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A
CC NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN
CC TRANSCRIPTIONAL LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING
CC INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
CC -!- SUBUNIT: FORMS HETERODIMER WITH TIMELESS (TIM); THE COMPLEX THEN

CC TRANSLOCATES INTO THE NUCLEUS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY.
CC FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE
CC TRANSLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED
CC FOR NUCLEAR LOCALIZATION (BY SIMILARITY).
CC -!- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY, PROBABLY BY THE
CC DOUBLE-TIME PROTEIN (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN
CC THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERODIMER
CC PER-TIM (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U11810; AAA76592.1; --
DR FlyBase; FBgn0015136; Dpic\per.
KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation.
FT NON TER 1
FT NON TER 143 143
SQ SEQUENCE 143 AA; 14459 MW; 5F4F2DC27F5D6846 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 143;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 AGPSLAA 29
| | | | |
DB 120 AGPSLAA 126

RESULT 19
Y244 METAC STANDARD; PRT; 144 AA.
AC Q8TU30;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein MA0244.
GN MA0244.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=1193238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirelli A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity."
RL Genome Res. 12:532-542(2002).
CC -!- SIMILARITY: BELONGS TO THE UPF0107 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE010682; AAC03697.1; -.
CC HAWAP; MF_00078; -. 1.
CC InterPro; IPR002840; DUF126.
CC Pfam; PF01989; DUF126; 1.
CC ProDom; PD015964; DUF126; 1.
CC KW Hypothetical protein; Complete proteome.
CC SEQUENCE 144 AA; 15343 MW; 54C250824262D44C CRC64;

Query Match 1.3%; Score 7; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 MLEKSPY 68
Db 108 MLEKSPY 114
|||||

RESULT 20
YS72 MYCTU STANDARD; PRT; 147 AA.
ID YS72 MYCTU
AC Q10800;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein RV2872.
GN RV2872 OR MT2939 OR MTCY274.03.
OS Mycobacterium tuberculosis.
CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
CC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CC NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Ruster S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z74024; CAA98372.1; -.
CC PIR; C70923; AAK47264.1; -.
CC TIGR; MT2939; -.
CC Tuberculist; RV2872; -.
CC InterPro; IPR006226; DUF_Myctu.
```

```
DR InterPro; IPR002716; PIN.
DR Pfam; PF01850; PIN; 1.
DR TIGRFAMs; TIGR00028; Mtu_fam_191; 1.
DR KW Hypothetical protein; Complete proteome.
DR SEQUENCE 147 AA; 16596 MW; F2DC73045F392553 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 LAAPAM 33
Db 74 LAAPAM 80
|||||

RESULT 21
YCS2 PORPU STANDARD; PRT; 174 AA.
ID YCS2 PORPU
AC P51192;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hypothetical 20.1 kDa protein ycf52 (ORF174).
GN YCF52.
OS Porphyra purpurea.
CC Chloroplast.
CC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
CC NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RA Reith M.E., Munnolland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RT genome.";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -1- SIMILARITY: BELONGS TO THE ACETYLTRANSFERASE FAMILY. YCF52
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U38804; AAC08078.1; -.
CC PIR; S73113; S73113.
CC InterPro; IPR000182; GCN5acetyltransf.
CC Pfam; PF00583; Acetyltransf; 1.
CC KW Hypothetical protein; Transferase; Acyltransferase; Chloroplast.
CC SEQUENCE 174 AA; 20099 MW; 8F8F7E09301D87C2 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 174;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 526 LIKOLRQ 532
Db 130 LIKOLRQ 136
|||||

RESULT 22
ALL7 DERPT STANDARD; PRT; 215 AA.
ID ALL7 DERPT
AC P49273;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mite allergen Der p 7 precursor (Der p VII).
GN DERP7.
OS Dermatophagoides pteronyssinus (House-dust mite).
CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
```

OC Acariformes; Sarcoptriformes; Astigmata; Psoroptidia; Analgoidea;
 OC Pyroglyphidae; Dermatophagoides.
 OX NCBI_TaxID=6956;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shen H.-D., Chua K.-Y., Lin K.-L., Heieh K.-H., Thomas W.R.;
 RT "Molecular cloning of a house dust mite allergen with common antibody
 binding specificities with multiple components in mite extracts.";
 RL Clin. Exp. Allergy 23:934-940(1993).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE MITE GROUP 7 ALLERGEN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U37044; AAA80264.1; -.
 KW Allergen; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 215 MITE ALLERGEN DER P 7.
 FT CARBOHYD 151 151 N-LINKED (GLCNAC...) (POTENTIAL).
 FT SEQUENCE 215 AA; 23881 MW; 63AD03DB5C8CB1C0 CRC64;
 SQ
 Query Match 1.3%; Score 7; DB 1; Length 215;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 400 MTLTSFE 406
 DB 152 MTLTSFE 158
 RESULT 23
 ID GL1A ARATH STANDARD; PRT; 223 AA.
 AC Q9FMA9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Germin-like protein subfamily 1 member 10 precursor.
 GN AT5G38930 OR K15E6.16 OR K15E6.110.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson B., Rohlfing T., Nelson J.,
 RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
 RA Martienssen R., McConbie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volktaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
 RA Entlan K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grynoprez B., Zimmermann W.,
 RA Ramspurger U., Wedler H., Balke K., Wedler E., Peters S.,

RA van Staveren M., Dirkee W., Moolijman P., Klein Lankhorst R.,
 RA Weitzneger T., Bothe G., Rose M., Hauf J., Berneiser S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K., Rudd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.,
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 thaliana";
 RT Nature 408:823-826(2000).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=98290546; PubMed=9628582;
 RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
 RT Sequence features of the regions of 1,456,315 bp covered by nineteen
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:41-54(1998).
 CC -!- FUNCTION: MAY PLAY A ROLE IN PLANT DEFENSE. HAS PROBABLY NO
 CC OXALATE OXIDASE ACTIVITY EVEN IF THE ACTIVE SITE IS CONSERVED.
 CC -!- SUBUNIT: OLIGOMER (BELIEVED TO BE A PENTAMER BUT PROBABLY
 CC HEXAMER) (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Apoplast (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE GERMIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB009048; BAB08649.1; -.
 DR HSSP; P45850; 1FI2.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin.sup.
 DR InterPro; IPR001929; Germin.
 DR Pfam; PF00190; Cupin; 1.
 DR PRINTS; PR00325; GERMIN.
 DR PROSITE; PS00725; GERMIN; 1.
 KW Apoplast; Cell wall; Signal; Glycoprotein; Manganese; Metal-binding;
 KW Multigene family.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 223 GERMIN-LIKE PROTEIN SUBFAMILY 1 MEMBER
 FT METAL 113 113 MANGANESE (BY SIMILARITY).
 FT METAL 115 115 MANGANESE (BY SIMILARITY).
 FT METAL 120 120 MANGANESE (BY SIMILARITY).
 FT METAL 161 161 MANGANESE (BY SIMILARITY).
 FT DISULFID 34 51 BY SIMILARITY.
 FT CARBOHYD 38 38 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 223 AA; 23631 MW; C24E679455B0BB8B CRC64;
 Query Match 1.3%; Score 7; DB 1; Length 223;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 169 TARPITS 175
 DB 69 TARPITS 75
 RESULT 24
 ID GL1B ARATH STANDARD; PRT; 223 AA.
 AC Q9FWA8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Germin-like protein subfamily 1 member 11 precursor.
 GN AT5G38940 OR K15E6.18 OR K15E6.120.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OS Arabidopsids thaliana (Mouse-ear cress).


```

SQ SEQUENCE 229 AA; 25867 MW; 5CBE57DBDB4C275 CRC64;
Query Match 1.3%; Score 7; DB 1; Length 229;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 NGQLFNN 387
DB 221 NGQLFNN 227

RESULT 26
VB08 VACCC
ID VB08 VACCC STANDARD; PRT; 272 AA.
AC P21004;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein B8 precursor.
GN B8R.
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10249;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021027; PubMed=2219722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paolletti E.;
RT "The complete DNA sequence of vaccinia virus.";
RL Virology 179:247-266(1990).
RN [2]
RP COMPLETE GENOME.
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paolletti E.;
RT "Appendix to 'The complete DNA sequence of vaccinia virus.'";
RL Virology 179:517-563(1990).
CC -!- SIMILARITY: TO SHOPE FIBROMA VIRUS T7 PROTEIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M35027; AAA48205.1; -.
DR PIR; G42526; G42526.
DR InterPro; IPR000282; Cytok receptor_2.
DR InterPro; IPR006987; Pox_IFNR.
DR Pfam; PF04903; Pox IFNR; 1.
KW Glycoprotein; Signal.
FT SIGNAL 1 13 POTENTIAL.
FT CHAIN 14 272 PROTEIN B8.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 272 AA; 31163 MW; 2680C169167D875 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 272;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 LNLTKYS 208
DB 266 LNLTKYS 272

RESULT 27
VB08 VACCV
ID VB08 VACCV STANDARD; PRT; 272 AA.
AC P24770;

```

```

DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Protein B8 precursor.
GN B8R.
OS Vaccinia virus (strain WR).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10254;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91259063; PubMed=2045793;
RA Smith G.L., Chan Y.S., Howard S.T.;
RT "Nucleotide sequence of 42 kbp of vaccinia virus strain WR from near
RL the right inverted terminal repeat.";
RN J. Gen. Virol. 72:1349-1376(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91111982; PubMed=1846491;
RA Howard S.T., Chan Y.S., Smith G.L.;
RT "Vaccinia virus homologues of the Shope fibroma virus inverted
RT terminal repeat proteins and a discontinuous ORF related to the tumor
RT necrosis factor receptor family.";
RL Virology 180:633-647(1991).
CC -!- SIMILARITY: TO SHOPE FIBROMA VIRUS T7 PROTEIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; D11079; BAA01838.1; -.
DR EMBL; M58056; AAA47967.1; -.
DR PIR; JQ1802; JQ1802.
DR InterPro; IPR000282; Cytok receptor_2.
DR InterPro; IPR006987; Pox_IFNR.
DR Pfam; PF04903; Pox IFNR; 1.
KW Glycoprotein; Signal.
FT SIGNAL 1 13 POTENTIAL.
FT CHAIN 14 272 PROTEIN B8.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 272 AA; 31158 MW; 3CTED994A82DDB93 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 272;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 LNLTKYS 208
DB 266 LNLTKYS 272

RESULT 28
VPM BPHPI
ID VPM BPHPI STANDARD; PRT; 281 AA.
AC P51721;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Probable terminase, endonuclease subunit (ORF19).
OS Bacteriophage HPI.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=10690;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HPIC1;
RX MEDLINE=96279738; PubMed=8710508;
RA Esposito D., Fitzmaurice W.P., Benjamin R.C., Goodman S.D.,

```


RA Waidman A.S., Scoocca J.J.;
 RT "The complete nucleotide sequence of bacteriophage HPI DNA.";
 RL Nucleic Acids Res. 24:2360-2368(1996).
 CC -!- SIMILARITY: TO PHAGE P2 PROTEIN M.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U24159; AAB09204.1; -;
 DR PIR; S69525; S69525.
 KW Hydrolase; Endonuclease; DNA-binding; Capsid assembly.
 SQ SEQUENCE 281 AA; 31834 MW; 2F21E3EC77052DFC CRC64;
 Query Match 1.3%; Score 7; DB 1; Length 281;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 407 RALSIAE 413
 Db 105 RALSIAE 111
 RESULT 29
 FMKB_ECOLI STANDARD; PRT; 285 AA.
 AC P02970;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, last sequence update)
 DT 01-JUN-1994 (Rel. 29, last annotation update)
 DE K88 fimbrial protein AB precursor (K88 pilin) (K88 antigen).
 GN FAGG.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Gastra W., Mooi F.R., Stuitje A.R., de Graaf F.K.;
 RT "The nucleotide sequence of the gene encoding the K88ab protein
 subunit of porcine enterotoxigenic Escherichia coli.";
 RL FEMS Microbiol. Lett. 12:41-46(1981).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86007039; PubMed=2412961;
 RA Dykes C.W., Halliday I.J., Read M.J., Hobden A.N., Harford S.;
 RT "Nucleotide sequences of four variants of the K88 gene of porcine
 enterotoxigenic Escherichia coli";
 RL Infect. Immun. 50:279-283(1985).
 RN (3)
 RP SEQUENCE OF 22-185.
 RC STRAIN=D1721;
 RX MEDLINE=82027186; PubMed=7026236;
 RA Klemm P.;
 RT "The complete amino-acid sequence of the K88 antigen, a fimbrial
 protein from Escherichia coli";
 RL Eur. J. Biochem. 117:617-627(1981).
 CC -!- FUNCTION: K88 MAJOR FIMBRIAL SUBUNIT. FIMBRIAE (ALSO CALLED PILI),
 CC POLAR FILAMENTS RADIATING FROM THE SURFACE OF THE BACTERIUM TO A
 CC LENGTH OF 0.5-1.5 MICROMETERS AND NUMBERING 100-300 PER CELL,
 CC ENABLE BACTERIA TO COLONIZE THE EPITHELIUM OF SPECIFIC HOST
 CC ORGANS.
 CC -!- SUBUNIT: K88 FIMBRIA, 0.1-1 MICROMETER IN LENGTH AND 7 NANOMETERS
 CC IN DIAMETER, IS COMPOSED OF ABOUT 100 IDENTICAL SUBUNITS.
 CC -!- MISCELLANEOUS: THE PROTEIN EXISTS IN SEVERAL ANTIGENIC VARIANTS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V00292; CAA23567.1; -;
 DR EMBL; M29374; AAA24032.1; -;
 DR PIR; S07208; Y0EC88.
 DR InterPro: IPR003467; Fimbrin K88.
 DR Pfam: PF02432; Fimbrin_K88; 1.
 KW Fimbria; Signal; Antigen; Plasmid.
 FT SIGNAL 1 21
 FT CHAIN 22 285 K88 FIMBRIAL PROTEIN AB.
 SQ SEQUENCE 285 AA; 29546 MW; 8ABBF6EC16616450 CRC64;
 Query Match 1.3%; Score 7; DB 1; Length 285;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGLTTTV 7
 Db 181 GGLTTTV 187
 RESULT 30
 BSN2_BACSU STANDARD; PRT; 288 AA.
 AC O32150;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Extracellular ribonuclease precursor (EC 3.1.-.-).
 GN BSN.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunz F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel P.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujić P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,
 RA Totsato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis".
 RL Nature 390:249-256(1997).
 CC -!- FUNCTION: MG(2+)-ACTIVATED RIBONUCLEASE WHICH HYDROLYZES RNA

CC APPARENTLY NONSPECIFICALLY INTO OLIGONUCLEOTIDES WITH 5'-TERMINAL
 CC PHOSPHATE (BY SIMILARITY).
 CC
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z99120; CAB15244.1; -;
 CC PIR; H70017; H70017;
 CC Subtilist; BG13995; bsn.
 CC Pfam; PF04231; Endonuclease I; 1.
 CC Hydrolase; Nuclease; Signal; Complete proteome.
 CC SIGNAL 1 26 POTENTIAL.
 CC CHAIN 27 288 EXTRACELLULAR RIBONUCLEASE.
 CC SEQUENCE 288 AA; 32094 MW; F884777DA27F8F7 CRC64;
 CC
 CC Query Match 1.3%; Score 7; DB 1; Length 288;
 CC Best Local Similarity 100.0%; Pred. No. 44;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 26 SLAAPAA 32
 CC DB |||||
 CC 50 SLAAPAA 56
 CC
 CC RESULT 31
 CC TF_RABIT
 CC ID TF_RABIT STANDARD; PRT; 292 AA.
 CC AC P24055;
 CC DT 01-MAR-1992 (Rel. 21, Created)
 CC DT 01-MAR-1992 (Rel. 21, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Tissue factor precursor (TF) (Coagulation factor III).
 CC F3.
 CC OS Oryctolagus cuniculus (Rabbit).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 CC OX NCBI_TaxID=9986;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX TISSUE=Heart;
 CC RC MEDLINE=91200676; PubMed=1840552;
 CC RA Andrews B.S., Rehmtulla A., Fowler B.J., Edgington T.S., Mackman N.;
 CC RT "Conservation of tissue factor primary sequence among three mammalian
 CC RT species."
 CC RL Gene 98:265-269(1991).
 CC RN [2]
 CC RP SEQUENCE OF 33-292 FROM N.A.
 CC RC STRAIN=New Zealand white; TISSUE=Brain;
 CC RX MEDLINE=92081032; PubMed=1746002;
 CC RA Pawashe A., Ezekowitz M., Lin T.C., Horton R., Bach R., Konigsberg W.;
 CC RT "Molecular cloning, characterization and expression of cDNA for
 CC RT rabbit brain tissue factor."
 CC RL Thromb. Haemost. 66:315-320(1991).
 CC RN [3]
 CC RP X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 36-240.
 CC RX MEDLINE=98266351; PubMed=9605315;
 CC RA Muller Y.A., Kelley R.F., de Vos A.M.;
 CC RT "Hinge bending within the cytokine receptor superfamily revealed by
 CC RT the 2.4 A crystal structure of the extracellular domain of rabbit
 CC RT tissue factor."
 CC RL Protein Sci. 7:1106-1115(1998).
 CC
 CC -!- FUNCTION: INITIATES BLOOD COAGULATION BY FORMING A COMPLEX WITH
 CC CIRCULATING FACTOR VII OR VIIA. THE [TF:VIIA] COMPLEX ACTIVATES
 CC FACTORS IX OR X BY SPECIFIC LIMITED PROLYSIS. TF PLAYS A ROLE IN
 CC NORMAL HEMOSTASIS BY INITIATING THE CELL-SURFACE ASSEMBLY AND
 CC PROPAGATION OF THE COAGULATION PROTEASE CASCADE.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC
 CC -!- TISSUE SPECIFICITY: BRAIN, HEART.
 CC -!- SIMILARITY: BELONGS TO THE TISSUE FACTOR FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M55390; AAA63469.1; -;
 CC PIR; X53521; CAA37597.1; -;
 CC PIR; JU0441; KFRB3.
 CC PDB; 1A21; 27-MAY-98.
 CC InterPro; IPR000282; Cytok_receptor_2.
 CC InterPro; IPR001187; Tissue_factor.
 CC Pfam; PF01108; Tissue_fac; 1.
 CC PRINTS; PR00346; TISSUEFACTOR.
 CC PROSITE; PS00621; TISSUE_FACTOR; 1.
 CC Glycoprotein; Blood coagulation; Transmembrane; Signal; Lipoprotein;
 CC Palmitate; 3D-structure.
 CC SIGNAL 1 32
 CC CHAIN 33 292 TISSUE FACTOR.
 CC DOMAIN 33 250 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 251 271 POTENTIAL.
 CC DOMAIN 272 292 CYTOPLASMIC (POTENTIAL).
 CC SITE 44 46 WKS MOTIF.
 CC SITE 75 77 WKS MOTIF.
 CC CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC DISULFID 79 87
 CC LIPID 216 239
 CC LIPID 274 274
 CC STRAND 40 47
 CC TURN 48 49
 CC STRAND 50 56
 CC STRAND 62 69
 CC STRAND 76 82
 CC STRAND 86 88
 CC HELIX 90 93
 CC TURN 94 95
 CC TURN 97 98
 CC STRAND 101 109
 CC STRAND 126 126
 CC STRAND 130 130
 CC STRAND 132 135
 CC STRAND 137 137
 CC STRAND 143 149
 CC TURN 150 151
 CC STRAND 152 157
 CC STRAND 161 165
 CC STRAND 170 172
 CC STRAND 173 177
 CC HELIX 178 180
 CC STRAND 182 188
 CC STRAND 196 200
 CC STRAND 204 208
 CC TURN 211 212
 CC STRAND 216 222
 CC TURN 224 225
 CC STRAND 231 231
 CC STRAND 238 239
 CC SEQUENCE 292 AA; 32738 MW; 4860A1CADBACCF71 CRC64;
 CC
 CC Query Match 1.3%; Score 7; DB 1; Length 292;
 CC Best Local Similarity 100.0%; Pred. No. 45;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 108 QVPRPGT 114

[illegible]

```
DE Putative homoserine O-acetyltransferase (EC 2.3.1.31) (Homoserine O-
DE trans-acetylase) (Homoserine transacetylase) (HTA).
CN MTH1820.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Delta H;
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + L-homoserine = CoA + O-acetyl-L-
CC homoserine.
CC -!- PATHWAY: Methionine biosynthesis; HTA variant; first step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the AB hydrolase superfamily. HTA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB000935; AAB86286.1; -.
DR PIR; D69110; D69110.
DR HAMAP; MF 00296; atypical; 1.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000379; Ser_estr_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PIRSF; PIRSF000443; Homoser Ac trans; 1.
KW Hypothetical protein; Methionine biosynthesis; Transferase;
KW Acyltransferase; Complete proteome.
FT ACT_SITE 134 134 POTENTIAL.
FT ACT_SITE 299 299 POTENTIAL.
SQ SEQUENCE 319 AA; 35687 MW; 46472BDE6F981F83 CRC64;
Query Match 1.3%; Score 7; DB 1; Length 319;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 406 ERALSIA 412
Db 191 ERALSIA 197
RESULT 35
PE30 ARATH
ID PE30 ARATH STANDARD; PRT; 329 AA.
AC Q9LSY7; Q43737; Q96521;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Peroxidase 30 precursor (EC 1.11.1.7) (Asperox P30) (PRXR5) (ATP7a).
GN PER30 OR P30 OR AT3G21770 OR MSD21.8 OR MSD21.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurobids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
```

```
RC STRAIN=cv. Columbia;
RA Capelli N., Tognolli M., Flach J., Overney S., Penel C., Greppin H.,
RA Simon P.;
RT "Eleven cDNA clones from Arabidopsis thaliana encoding
RT isoperoxidases.";
RL (In) Plant Gene Register PGR96-066.
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RC STRAIN=cv. Columbia;
RA Welinder K.G., Jespersen H.M., Kjaersgaard I.V.H., Justesen A.F.,
RA Oestergaard L., Abelskov A.K., Jensen L.N.,
RA Rasmussen S.K.;
RT "From expressed sequence tags to structure, function, evolution and
RT expression of 28 ER-targeted Arabidopsis peroxidases.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RC STRAIN=cv. Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RA Braver V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RC STRAIN=cv. Columbia;
RA Shinzaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (ARLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGECC).";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RN CHARACTERIZATION
RP STRAIN=cv. Columbia;
RC STRAIN=cv. Columbia;
RX MEDLINE=98409321; PubMed=9738941;
RA Oestergaard L., Pedersen A.G., Jespersen H.M., Brunak S.,
RA Welinder K.G.;
RT "Computational analyses and annotations of the Arabidopsis peroxidase
RT gene family.";
RL FEBS Lett. 433:98-102(1998).
RN [7]
RN TISSUE SPECIFICITY.
RP STRAIN=cv. Columbia;
RC STRAIN=cv. Columbia;
RX MEDLINE=99024953; PubMed=9807821;
RA Ruan Y., Gilmore J., Conner T.;
RT "Towards Arabidopsis genome analysis: monitoring expression profiles
RT of 1400 genes using cDNA microarrays.";
RL Plant J. 15:821-833(1998).
RN [8]
RN TISSUE SPECIFICITY.
RP STRAIN=cv. Columbia;
RC STRAIN=cv. Columbia;
RA Zhu T., Sudworth P., Han B., Brown D., Chang H.-S., Zou G., Wang X.;
RT "Toward elucidating the global gene expression patterns of developing
RT Arabidopsis: parallel analysis of 8300 genes by a high-density
RT oligonucleotide probe array.";
RL Plant Physiol. Biochem. 39:221-242(2001).
RN [9]
RN GENE FAMILY ORGANIZATION, AND NOMENCLATURE.
RP STRAIN=cv. Columbia;
RC STRAIN=cv. Columbia;
RX MEDLINE=22030461; PubMed=12034502;
RA Tognolli M., Penel C., Greppin H., Simon P.;
RT "Analysis and expression of the class III peroxidase large gene family
RT in Arabidopsis thaliana.";
RL Gene 288:129-138(2002).
CC -!- FUNCTION: Removal of H(2)O(2), oxidation of toxic reductants,
CC biosynthesis and degradation of lignin, suberization, auxin
CC catabolism, response to environmental stresses such as wounding,
CC
```

pathogen attack and oxidative stress. These functions might be dependent on each isozyme/isoform in each plant tissue.

-1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.

-1- COFACTOR: Binds 1 prothemo IX, 1 iron(III) ion and 2 calcium ions.

-1- SUBCELLULAR LOCATION: Secreted (By similarity).

-1- TISSUE SPECIFICITY: Mainly expressed in roots.

-1- INDUCTION: Induced by high salinity stress.

-1- MISCELLANEOUS: There are 73 peroxidase genes in A.thaliana.

-1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. CLASSICAL PLANT (CLASS III) PEROXIDASE SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X98321; CAA66965.1; -;
 EMBL; X98854; CAA67360.1; ALT_INIT.
 EMBL; AB025634; BAB02839.1; -;
 EMBL; AY084816; AAM61382.1; ALT_INIT.
 EMBL; AY072326; AAL61933.1; -;
 EMBL; AY114567; AAM47886.1; -;
 HSSP; P00433; 3ATJ
 IncerPro; IPR002016; Peroxidase.
 Pfam; PF00141; peroxidase; 1.
 PRINTS; PR00458; PEROXIDASE.
 PROSITE; PS00435; PEROXIDASE 1; 1.
 PROSITE; PS00436; PEROXIDASE 2; 1.
 PROSITE; PS00873; PEROXIDASE 4; 1.
 Oxidoreductase; Glycoprotein; Peroxidase; Iron; Heme;
 Multigene family; Calcium; Signal; Pyrrolidone carboxylic acid.
 SIGNAL 1 27 POTENTIAL.
 CHAIN 28 329 PEROXIDASE 30.
 METAL 70 70 CALCIUM 1 (BY SIMILARITY).
 METAL 73 73 CALCIUM 1 (VIA CARBONYL OXYGEN)
 (BY SIMILARITY).
 METAL 75 75 CALCIUM 1 (VIA CARBONYL OXYGEN)
 (BY SIMILARITY).
 METAL 77 77 CALCIUM 1 (BY SIMILARITY).
 METAL 79 79 CALCIUM 1 (BY SIMILARITY).
 METAL 195 195 CALCIUM 2 (BY SIMILARITY).
 METAL 247 247 CALCIUM 2 (BY SIMILARITY).
 METAL 250 250 CALCIUM 2 (BY SIMILARITY).
 METAL 255 255 CALCIUM 2 (BY SIMILARITY).
 ACT_SITE 65 65 BY SIMILARITY.
 ACT_SITE 69 69 DISTAL HISTIDINE.
 ACT_SITE 96 96 HYDROGEN-BOUND (BY SIMILARITY).
 ACT_SITE 164 164 SUBSTRATE BINDING (BY SIMILARITY).
 METAL 194 194 IRON (HEME AXIAL LIGAND).
 ACT_SITE 272 272 HYDROGEN-BOUND (BY SIMILARITY).
 DISULFID 38 116 BY SIMILARITY.
 DISULFID 71 76 BY SIMILARITY.
 DISULFID 122 324 BY SIMILARITY.
 DISULFID 201 234 BY SIMILARITY.
 MOD_RES 28 28 PYRROLIDONE CARBOXYLIC ACID
 (BY SIMILARITY).
 CARBOHYD 83 83 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 169 169 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 210 210 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 240 240 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
 CONFLICT 133 133 A -> T (IN REF. 4).
 CONFLICT 156 157 KT -> LK (IN REF. 4).
 CONFLICT 174 174 Q -> R (IN REF. 4).
 CONFLICT 329 329 S -> N (IN REF. 1 AND 2).
 SEQUENCE 329 AA; 35788 MW; 78CE941617393A0 CRC64;

Query Match

1.3%; Score 7; DB 1; Length 329;

Best Local Similarity 100.0%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 GPSLAAP 30
 DB 56 GPSLAAP 62
 |||||

RESULT 36
 Y258 HAEIN
 ID Y258 HAEIN STANDARD; PRT; 330 AA.
 AC P43974; P44597; P71347;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein HI0258.
 GN HI0258.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-P., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.";
 RL Science 269:496-512 (1995).
 RN [2]
 RP REVISIONS.
 RA White O., Clayton R.A., Kerlavage A.R., Fleischmann R.D.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO N.GONORRHOEA LGTIC AND WEAK, TO B.SUBTILIS GSPA,
 CC LIPOPOLYSACCHARIDE 1,3-GALACTOSYLTRANSFERASE (RFAI) AND TO
 CC LIPOPOLYSACCHARIDE 1,2-GLUCOSYLTRANSFERASE (RFAJ).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; U32711; AAC21924.1; -;
 TIGR; HI0258; -;
 DR InterPro; IPR002495; Glyco trans 8.
 DR Pfam; PF01501; Glyco_transf_8; 1-
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 330 AA; 39073 MW; C5A6CA0207F24P9A CRC64;

Query Match 1.3%; Score 7; DB 1; Length 330;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 RLNLTKY 207
 DB 123 RLNLTKY 129
 |||||

RESULT 37
 RFCS HUMAN
 ID RFCS HUMAN STANDARD; PRT; 340 AA.
 AC P40937;

DR PIR; A40153; LUPY1.
 DR HSP; P19619; IHM6.
 DR InterPro: IPR001464; Annexin.
 DR Pfam; PF00191; annexin; 4.
 DR PRINTS; PR00196; ANNEXIN.
 DR PRODOM; PD000143; Annexin; 4.
 DR SMART; SM00335; ANX; 4.
 DR PROSITE; PS00223; ANNEXIN; 3.
 KW Annexin: Calcium/phospholipid-binding; Repeat;
 KW Phospholipase A2 inhibitor.
 FT REPEAT 46 106 ANNEXIN 1.
 FT REPEAT 118 178 ANNEXIN 2.
 FT REPEAT 202 262 ANNEXIN 3.
 FT REPEAT 276 336 ANNEXIN 4.
 SQ SEQUENCE 341 AA; 38451 MW; AB6B5CF7F3618DA2 CRC64;
 Query Match 1.3%; Score 7; DB 1; Length 341;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 336 YYKEVLK 342
 DB 150 YYKEVLK 156
 |||||
 RESULT 39
 AN12 COLLI
 ID AN12 COLLI STANDARD; PRT; 343 AA.
 AC Q92040; Q92041;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Annexin I, isoform P37 (Lipocortin I) (Calpactin II) (Chromobindin 9)
 DE (Phospholipase A2 inhibitory protein).
 GN CP37.
 OS Columba livia (Domestic pigeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
 OX NCBI_TaxID=8932;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 13-50.
 RC STRAIN=White Carneau; TISSUE=Cropac;
 RX MEDLINE=92406850; PubMed=1388165;
 RA Haigler H.T., Mangili J.A., Gao Y., Jones J., Horeman N.D.;
 RT "Identification and characterization of columbid annexin Icp37.
 RT Insights into the evolution of annexin I phosphorylation sites.";
 RL J. Biol. Chem. 267:19123-19129(1992).
 RN [2]
 RP SEQUENCE OF 1-154 FROM N.A.
 RC STRAIN=N41; TISSUE=Liver;
 RX MEDLINE=94266150; PubMed=8206371;
 RA Gao Y., Horeman N.D.;
 RT "Structural and functional divergences of the columbid annexin I-
 RT encoding cp37 and cp35 genes.";
 RL Gene 143:179-186(1994).
 CC -1- FUNCTION: CALCIUM/PHOSPHOLIPID-BINDING PROTEIN WHICH PROMOTES
 CC MEMBRANE FUSION AND IS INVOLVED IN EXOCYTOSIS. THIS PROTEIN
 CC REGULATES PHOSPHOLIPASE A2 ACTIVITY. IT SEEMS TO BIND FROM TWO TO
 CC FOUR CALCIUM IONS WITH HIGH AFFINITY.
 CC -1- DOMAIN: A pair of annexin repeats may form one binding site for
 CC calcium and phospholipid.
 CC -1- PTM: PHOSPHORYLATED BY PROTEIN KINASE C AND EPIDERMAL GROWTH
 CC FACTOR RECEPTOR/KINASE.
 CC -1- PTM: THE N-TERMINUS IS BLOCKED.
 CC -1- MISCELLANEOUS: IN PIGEONS, TWO ISOFORMS OF ANNEXIN-I ARE ENCODED
 CC BY THE DIFFERENTIALLY REGULATED GENES CP35 AND CP37.
 CC -1- SIMILARITY: BELONGS TO THE ANNEXIN FAMILY.
 CC -1- SIMILARITY: Contains 4 annexin repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M91008; AAA49447.1; ALT_INIT.
 DR EMBL; L02504; AAA20674.1; -.
 DR PIR; A4118; A4118.
 DR HSP; P19619; IHM6.
 DR InterPro: IPR001464; Annexin.
 DR Pfam; PF00191; annexin; 4.
 DR PRINTS; PR00196; ANNEXIN.
 DR PRODOM; PD000143; Annexin; 4.
 DR SMART; SM00335; ANX; 4.
 DR PROSITE; PS00223; ANNEXIN; 4.
 KW Annexin: Calcium/phospholipid-binding; Repeat;
 KW Phospholipase A2 inhibitor; Phosphorylation.
 FT REPEAT 47 107 ANNEXIN 1.
 FT REPEAT 119 179 ANNEXIN 2.
 FT REPEAT 203 264 ANNEXIN 3.
 FT REPEAT 278 338 ANNEXIN 4.
 FT MOD_RES 19 19 INTERCHAIN CROSS-LINK (BY SIMILARITY).
 FT MOD_RES 21 21 PHOSPHORYLATION (BY EGR).
 FT MOD_RES 24 24 PHOSPHORYLATION (BY PKC).
 FT CONFLICT 74 75 HR -> QQ (IN REF. 2).
 SQ SEQUENCE 343 AA; 38849 MW; 49A621A92BCEDFDC CRC64;
 Query Match 1.3%; Score 7; DB 1; Length 343;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 336 YYKEVLK 342
 DB 151 YYKEVLK 157
 |||||
 RESULT 40
 YES2 LISMO
 ID YES2 LISMO STANDARD; PRT; 373 AA.
 AC P53434;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein Lmo1452.
 GN Lmo1452.
 OS Listeria monocytogenes
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LO28 / Serovar 1/2c;
 RA Klarsfeld A.D., Cossart P.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EGD-e / Serovar 1/2a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Feihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
 RA Madueno E., Maicouram A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomics of Listeria species.";
 RL Science 294:849-852(2001).
 CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U17284; AAA62502.1; --
 DR EMBL; AL591979; CAC99530.1; --
 DR PIR; AD1256; AD1256.
 DR ListList; LMO01452; --
 DR InterPro; IPR002678; DUF34.
 DR Pfam; PF01784; NIF3; 1.
 DR TIGRFAMs; TIGR00486; TIGR00486; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 373 AA; 41433 MW; C79AF8A87E22369A2 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 373;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 443 RLALVNN 449
 |||||
 Db 152 RLALVNN 158

RESULT 41

ID YE89 LISIN STANDARD; PRT; 373 AA.
 AC Q92B08;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein Lin1489.
 GN LIN1489.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN-CLIP 11262 / Serovar 6a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetoui F., Couve E., de Daruvar A., Deloux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapat G.,
 RA Madieno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
 RT "Comparative genomics of *Listeria species*."
 RL Science 294:849-852(2001).
 CC -!- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AL596168; CAC96720.1; --
 DR PIR; AH1618; AH1618.
 DR ListList; LIN01489; --
 DR InterPro; IPR002678; DUF34.
 DR Pfam; PF01784; NIF3; 1.
 DR TIGRFAMs; TIGR00486; TIGR00486; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 373 AA; 41436 MW; 4B8C6A93B0B9F540 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 373;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 443 RLALVNN 449
 |||||
 Db 152 RLALVNN 158

RESULT 42

ID ALFC CHLRE STANDARD; PRT; 374 AA.
 AC Q42690; Q36725;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Fructose-bisphosphate aldolase 1, chloroplast precursor (EC 4.1.2.13).
 GN ALBCHL.
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadales; Chlamydomonadales.
 OX NCBI_TaxID=3055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94330714; PubMed=8053679;
 RA Schnarrenberger C., Pelzer-Reith B., Yatsuki H., Freund S.,
 RA Jacobshagen S., Hori K.;
 RT "Expression and sequence of the only detectable aldolase in
 RT Chlamydomonas reinhardtii."
 RL Arch. Biochem. Biophys. 313:173-178(1994).
 CC -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
 CC phosphate + D-glyceraldehyde 3-phosphate.
 CC -!- PATHWAY: Fructosebisphosphate and sedoheptulosebisphosphate
 CC biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast.
 CC -!- SIMILARITY: Belongs to class I fructose-bisphosphate aldolase
 CC family.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X69969; CAA49590.1; --
 DR EMBL; S72951; AAC60574.1; --
 DR PIR; S48639; S48639.
 DR HSP; P14223; IAS5.
 DR InterPro; IPR000741; Aldolase_I.
 DR Pfam; PF00274; glycolytic_enz; 1.
 DR ProDom; PD001128; Aldolase_I; 1.
 DR PROSITE; PS00158; ALDOLASE_CLASS_I; FALSE NEG.
 KW Lyase; Schiff base; Glycolysis; Chloroplast; Transit peptide.
 FT TRANSIT 1 27 CHLOROPLAST (POTENTIAL).
 FT CHAIN 28 374 FRUCTOSE-BISPHOSPHATE ALDOLASE 1.
 FT BINDING 74 74 C-1-PHOSPHATE GROUP OF THE SUBSTRATE
 FT BINDING 162 162 (BY SIMILARITY).
 FT BINDING 162 162 C-1-PHOSPHATE GROUP OF THE SUBSTRATE
 FT BINDING 243 243 (BY SIMILARITY).
 FT BINDING 243 243 SCHIFF-BASE WITH DIHYDROXYACETONE-P
 FT ACT_SITE 374 374 ESSENTIAL FOR ENHANCED ACTIVITY OF THE
 FT ENZYME TOWARD FRUCTOSE 1,6-BISPHOSPHATE
 FT AS COMPARED WITH FRUCTOSE 1-PHOSPHATE
 FT (BY SIMILARITY).
 SQ SEQUENCE 374 AA; 40292 MW; B96340FDS5C2364B CRC64;

Query Match 1.3%; Score 7; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

QY 332 SAAEYK 338
    |||||
Db 149 SAAEYK 155

RESULT 43
GLF1_KLEPN STANDARD; PRT; 384 AA.
AC Q4845;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Probable UDP-galactopyranose mutase (EC 5.4.99.9).
GN RPB.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K20 / Serotype O1;
RX MEDLINE=9603272; PubMed=759323;
RA Clarke B.R., Bronner D., Keenleyside W.J., Severn W.B., Richards J.C.,
RA Whitefield C.;
RT "Role of Rfe and RfbF in the initiation of biosynthesis of D-galactan
RT I, the lipopolysaccharide O antigen from Klebsiella pneumoniae
RT serotype O1";
RL J. Bacteriol. 177:5411-5418 (1995).
CC -1- FUNCTION: INVOLVED IN THE CONVERSION OF UDP-GALP INTO UDP-GALF
CC THROUGH A 2-KETO INTERMEDIATE.
CC -1- CATALYTIC ACTIVITY: UDP-D-galactopyranose = UDP-D-galacto-1,4-
CC furanose.
CC -1- COPACTOR: FAD (BY SIMILARITY).
CC -1- PATHWAY: Lipopolysaccharide O antigen biosynthesis.
CC -1- SIMILARITY: N-TERMINAL SHOWS SIMILARITY TO FAD OR NAD CONTAINING
CC PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L31762; AAC98417.1; -.
DR InterPro; IPR00205; NAD binding.
DR InterPro; IPR004379; UDP-GALP_mutase.
DR Pfam; PF03275; GLF; 1.
DR TIGRFAMs; TIGR00031; UDP-GALP mutase; 1.
KW Lipopolysaccharide biosynthesis; Isomerase; Flavoprotein; FAD.
FT NP_BIND 5 33 FAD (POTENTIAL).
SQ SEQUENCE 384 AA; 44457 MW; B74EC65EA5B751AB CRC64;

Query Match 1.3%; Score 7; DB 1; Length 384;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 409 LSLAENE 415
    |||||
Db 329 LSLAENE 335

RESULT 44
BIOF_ERWHE STANDARD; PRT; 385 AA.
AC Q47829;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 8-amino-7-oxononanoate synthase (EC 2.3.1.47) (AONS) (8-amino-7-
DE ketopelargonate synthase) (7-keto-8-amino-pelargonate acid synthetase)
DE (7-KAP synthetase) (L-alanine--pimelyl CoA ligase).

GN BIOF.
OS Erwinia herbicola.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pantoea.
OX NCBI_TaxID=549;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EHO10;
RX MEDLINE=97217600; PubMed=9063571;
RA Wu C.H., Bao Y.Y., Shao C.P., Shiuan D.;
RT "Molecular cloning and nucleotide sequencing of bioF (7-keto-8-amino
RT pelargonate acid synthetase), bioC and bioD (dethiobiotin synthetase)
RT genes of Erwinia herbicola.";
RL Biochem. Mol. Biol. Int. 41:311-315 (1997).
CC -1- CATALYTIC ACTIVITY: 6-carboxyhexanoyl-CoA + L-alanine = 8-amino-7-
CC oxononanoate + CoA + CO(2).
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- PATHWAY: Biotin biosynthesis; first step.
CC -1- SIMILARITY: BELONGS TO CLASS-II OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U50183; BAB03272.1; -.
DR HSP; P12998; IBSO.
DR InterPro; IPR003408; Ala synthase.
DR InterPro; IPR004839; Aminotransf1/2.
DR InterPro; IPR004723; BioF.
DR InterPro; IPR001917; Nhttransf_2.
DR Pfam; PF02490; Ala_synthase; 1.
DR Pfam; PF00155; aminotran_1_2; 1.
DR TIGRFAMs; TIGR00858; bioF; 1.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
KW Biotin biosynthesis; Transferase; Pyridoxal phosphate.
FT BINDING 236 236 PYRIDOXAL PHOSPHATE (PROBABLE).
SQ SEQUENCE 385 AA; 41276 MW; BC30F6432DCE6B2 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 385;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 407 RALSALAE 413
    |||||
Db 330 RALSALAE 336

RESULT 45
TRBL_AGRU STANDARD; PRT; 395 AA.
ID TRBL_AGRU
AC PS4913;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Conjugal transfer protein trbl.
GN TRBL.
OS Agrobacterium tumefaciens.
OG Plasmid pTiA6NC.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=358;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312368; PubMed=8763954;
RA Alt-Morbe J., Stryker J.L., Fuqua C., Li P.L., Farrand S.K.,
RA Winans S.C.;
RT "The conjugal transfer system of Agrobacterium tumefaciens
RT octopine-type Ti plasmids is closely related to the transfer system

```

RT of an IncP plasmid and distantly related to Ti plasmid vir genes.";
RL J. Bacteriol. 178:4248-4257(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF242881; AAB95100.1; --
DR Pfam; PF04610; TrbL; 1.
KW Conjugation; Plasmid; Transmembrane.
FT TRANSMEM 9 29 POTENTIAL.
FT TRANSMEM 59 79 POTENTIAL.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 162 182 POTENTIAL.
FT TRANSMEM 183 203 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 255 275 POTENTIAL.
FT TRANSMEM 302 322 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT DOMAIN 166 170 POLY-VAL.
FT DOMAIN 385 390 POLY-PRO.
SQ SEQUENCE 395 AA; 40042 MW; 650F5F40936922D0 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 395;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 QSGRPGT 156
Db 378 QSGRPGT 384

RESULT 46
ID -SYW_RALSO STANDARD; PRT; 400 AA.
AC Q8Y0A1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)
DE (TrpS).
GN TRPS OR RSC1143 OR RS04610.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21681879; PubMed=11823852;
STRAIN=GMI1000;
RA Salanoubat M., Genin S., Attiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choiane N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL646062; CAD14845.1; --
DR HAMAP; MF 00140; atypical; 1.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; tRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGREMS; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT DOMAIN 173 241 INSERT.
FT SITE 12 20 "HIGH" REGION.
FT SITE 265 269 "KMSKS" REGION.
FT BINDING 268 268 ATP (BY SIMILARITY).
SQ SEQUENCE 400 AA; 45484 MW; D551F6CE02657DAA CRC64;

Query Match 1.3%; Score 7; DB 1; Length 400;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 470 EQRALL 476
Db 213 EQRALL 219

RESULT 47
CLPX_AQUAE STANDARD; PRT; 412 AA.
ID CLPX_AQUAE
AC O67356;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP-dependent Clp protease ATP-binding subunit clpX.
GN CLPX OR AQ_1337.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Auja M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC It directs the protease to specific substrates. Can perform
CC chaperone functions in the absence of clpP (By similarity).
CC -!- SUBUNIT: Heterodimer of clpP and clpX (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000735; AAC07316.1; --
DR PIR; A70416; A70416.
DR HSP; P32168; 1E94.
DR HAMAP; MF 00175; --; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR004487; ClpX.

```
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00382; ClpX; 1.
KW Chaperone; ATP-binding; Zinc-finger; Complete proteome.
FT ZN FING 10 36
FT NP BIND 117 124 C4-TYPE.
SQ SEQUENCE 412 AA; 46140 MW; PFCEAF451A930B69 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 VYIDEID 90
Db 179 VYIDEID 185

RESULT 48
CLPX_NEIMA STANDARD; PRT; 414 AA.
AC Q9JTX8;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP-dependent Clp protease ATP-binding subunit clpX.
GN CLPX OR NWA1585
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4B;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Davies S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis 22491."
RL Nature 404:502-506(2000).
CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC It directs the protease to specific substrates. Can perform
CC chaperone functions in the absence of clpp (By similarity).
CC -!- SUBUNIT: Heterodimer of clpp and clpX (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL162756; CAB84812.1; -
CC PIR; D81851; D81851.
CC HSP; P32168; 1E94.
CC HAMAP; MF 00175; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003959; AAA_ATPase_cent.
CC Pfam; PF00004; AAA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR00382; clpX; 1.
KW Chaperone; ATP-binding; Zinc-finger; Complete proteome.
FT ZN FING 8 33 C4-TYPE.
FT NP BIND 119 126 ATP (POTENTIAL).
SQ SEQUENCE 414 AA; 45198 MW; A3E417358F9F2981 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 VYIDEID 90
Db 181 VYIDEID 187

RESULT 49
CLPX_NEIMB STANDARD; PRT; 414 AA.
AC Q9JYJ3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP-dependent Clp protease ATP-binding subunit clpX.
GN CLPX OR NMB1372.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Ciftone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."
RL Science 287:1809-1815(2000).
CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC It directs the protease to specific substrates. Can perform
CC chaperone functions in the absence of clpp (By similarity).
CC -!- SUBUNIT: Heterodimer of clpp and clpX (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002485; AAP41746.1; -
CC PIR; A81091; A81091.
CC HSP; P32168; 1E94.
CC TIGR; NMB1372; -
CC HAMAP; MF 00175; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003959; AAA_ATPase_cent.
CC Pfam; PF00004; AAA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR00382; clpX; 1.
KW Chaperone; ATP-binding; Zinc-finger; Complete proteome.
FT ZN FING 8 33 C4-TYPE.
FT NP BIND 119 126 ATP (POTENTIAL).
SQ SEQUENCE 414 AA; 45099 MW; BC48D54586E085F1 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 VYIDEID 90
Db 181 VYIDEID 187
```

```

Db 215 ERALSIA 221

RESULT 50
HEMI_VIBCH STANDARD; PRT; 419 AA.
ID HEMI_VIBCH STANDARD; PRT; 419 AA.
AC Q9K024;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.-) (GLUTR).
GN HEMA OR VC2180.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bengal;
RA Ravichandran M., Lalitha P., Tang T.H., Chan Y.Y., Johari M.R.,
RA Zainuddin Z.F.;
RT "Cloning and characterization of hemA and hemM gene of Vibrio cholerae
RT (Bengal strain).";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterback T., Fleischmann R.D., Niernan W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483 (2000)
CC -!- CATALYTIC ACTIVITY: Glutamyl-tRNA (Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + tRNA(Glu).
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC -!- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF227752; AAK00701.1; -.
CC EMBL; AE004289; AAF95325.1; -.
CC PIR; C82109; C82109.
CC HSP; Q42843; 1B29.
CC TIGR; VC2180; -.
CC HAMAP; MF_00087; -.
CC InterPro; IPR000343; GLUTR.
CC Pfam; PF00745; GLUTR_dimer; 1.
CC Pfam; PF05201; GLUTR_N; 1.
CC Pfam; PF05200; GLUTR_NAD_bind; 1.
CC TIGRfam; TIGR01035; hemA; 1.
CC PROSITE; PS00747; GLUTR; 1.
CC Porphyrin biosynthesis; Oxidoreductase; NADP; Complete proteome.
KW ACT SITE 50 50 NUCLEOPHILE (BY SIMILARITY).
FT ACT SITE 97 97 BASE (BY SIMILARITY).
SQ SEQUENCE 419 AA; 45744 MW; 6D2D870AD3EEF8EC CRC64;

Query Match 1.3%; Score 7; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 406 ERALSIA 412
|||||

```

```

Db 215 ERALSIA 221

RESULT 51
CLPX_CAUCR STANDARD; PRT; 420 AA.
ID CLPX_CAUCR STANDARD; PRT; 420 AA.
AC O87708;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP-dependent Clp protease ATP-binding subunit clpX.
GN CLPX OR CC1961.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CB15N / NA1000;
RX MEDLINE=9255527; PubMed=10322004;
RA Osterae M., Stotz A., Schmid Nuoffer S., Jenal U.;
RT "Identification and transcriptional control of the genes encoding the
RT Caulobacter crescentus ClpXP protease.";
RL J. Bacteriol. 181:3039-3050 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Niernan W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Debocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uitterback T., Tran K., Wolf A., Vamathevan J., Fraser C.M.;
RA Salzberg S.L., Venter J.C., Shapiro L., Ermolaeva M., White O.,
RA "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC It directs the protease to specific substrates. Can perform
CC chaperone functions in the absence of clpP (By similarity).
CC -!- SUBUNIT: Heterodimer of clpP and clpX (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ010321; CAA09092.1; -.
CC EMBL; AE005869; AAK23936.1; -.
CC PIR; D87492; D87492.
CC HSP; P32168; 1E94.
CC TIGR; CC1961; -.
CC HAMAP; MF_00175; -.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003959; AAA_ATPase_cent.
CC InterPro; IPR004487; ClpX.
CC Pfam; PF00004; AAA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRfam; TIGR00382; clpX; 1.
CC Chaperone; ATP-binding; Zinc-finger; Complete proteome.
KW ZN FING 15 40 C4-TYPE.
FT NP BIND 118 125 ATP (POTENTIAL).
SQ SEQUENCE 420 AA; 45859 MW; 036339E2AB315C11 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 420;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VVIDEID 90

```

```
Db      180 VYIDEID 186
|||||||
RESULT 52
CLPX_ECOLI
ID      CLPX_ECOLI      STANDARD;      PRT;      423 AA.
AC      P33138;
DT      01-OCT-1993 (Rel. 27, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      ATP-dependent Clp protease ATP-binding subunit clpX.
GN      CLPX OR LOC OR B0438 OR C0554 OR Z0543 OR ECS0492 OR SF0383.
OS      Escherichia coli.
OS      Escherichia coli O6.
OS      Escherichia coli O157:H7, and
OS      Shigella flexneri.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Escherichia.
OX      NCBI_TaxID=562, 217992, 83334, 623;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      SPECIES=E.coli; STRAIN=K12;
RX      MEDLINE=94043020; PubMed=8226770;
RA      Gottesman S., Clark W.P., de Crecy-Lagard V., Maurizi M.R.;
RT      "ClpX, an alternative subunit for the ATP-dependent Clp protease of
RT      Escherichia coli. Sequence and in vivo activities.";
RL      J. Biol. Chem. 268:22618-22626(1993).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      SPECIES=E.coli;
RA      Yoo S., Seol J., Ha D., Goldberg A., Chung C.;
RL      Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RC      SPECIES=E.coli; STRAIN=K12 / MG1655;
RX      MEDLINE=97426617; PubMed=9278503;
RA      Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12.";
RL      Science 277:1453-1474(1997).
RN      [4]
RP      SEQUENCE FROM N.A.
RC      SPECIES=E.coli;
RA      Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
RA      Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA      Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;
RL      Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      SEQUENCE FROM N.A.
RC      SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX      MEDLINE=22388234; PubMed=12471157;
RA      Welch R.A., Burland V., Plunkett G. III, Redford P., Roesech P.,
RA      Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA      Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA      Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT      "Extensive mosaic structure revealed by the complete genome sequence
RT      of uropathogenic Escherichia coli.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RN      [6]
RP      SEQUENCE FROM N.A.
RC      SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX      MEDLINE=21074935; PubMed=11206551;
RA      Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA      Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA      Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA      Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA      Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA      Welch R.A., Blattner F.R.;
RT      "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL      Nature 409:529-533(2001).
RN      [7]
RP      SEQUENCE FROM N.A.
RC      SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX      MEDLINE=21156231; PubMed=11258796;
RA      Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA      Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA      Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA      Kuhara S., Shiba T., Hattori M., Shingawa H.;
RT      "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT      O157:H7 and genomic comparison with a laboratory strain K-12.";
RL      DNA Res. 8:11-22(2001).
RN      [8]
RP      SEQUENCE OF 1-24, AND CHARACTERIZATION.
RC      SPECIES=E.coli; STRAIN=K12 / W3110;
RX      MEDLINE=94043019; PubMed=8226769;
RA      Wojtkowiak D., Georgopoulos C., Zylicz M.;
RT      "Isolation and characterization of ClpX, a new ATP-dependent
RT      specificity component of the Clp protease of Escherichia coli.";
RL      J. Biol. Chem. 268:22609-22617(1993).
RN      [9]
RP      SEQUENCE FROM N.A.
RC      SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX      MEDLINE=22272406; PubMed=12384590;
RA      Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA      Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA      Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA      Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA      Yu J.;
RT      "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT      through comparison with genomes of Escherichia coli K12 and O157.";
RL      Nucleic Acids Res. 30:4432-4441(2002).
RN      [10]
RP      CHARACTERIZATION.
RC      SPECIES=E.coli;
RX      MEDLINE=95262627; PubMed=7743994;
RA      Wawrzynow A., Wojtkowiak D., Marszalek J., Banecki B., Jonsen M.,
RA      Graves B., Georgopoulos C., Zylicz M.;
RT      "The ClpX heat-shock protein of Escherichia coli, the ATP-dependent
RT      substrate specificity component of the ClpX-ClpX protease, is a novel
RT      molecular chaperone.";
RL      EMBO J. 14:1867-1877(1995).
RN      [11]
RP      FUNCTION: ATP-DEPENDENT SPECIFICITY COMPONENT OF THE CLP PROTEASE.
CC      IT DIRECTS THE PROTEASE TO SPECIFIC SUBSTRATES. IT MAY BIND TO THE
CC      LAMBDA O SUBSTRATE PROTEIN AND PRESENT IT TO THE CLP PROTEASE IN
CC      A FORM THAT CAN BE RECOGNIZED AND READILY HYDROLYZED BY CLPP. CAN
CC      PERFORM CHAPERONE FUNCTIONS IN THE ABSENCE OF CLPP.
CC      -!- SUBUNIT: HETERODIMER OF CLPP AND CLPX.
CC      -!- INDUCTION: By heat shock.
CC      -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
RN      [12]
RP      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
RN      [13]
RP      EMBL; L18867; AAA16116.1; -
DR      EMBL; Z23278; CAA80816.1; -
DR      EMBL; AE000150; AAC73541.1; -
DR      EMBL; U82664; BAB40194.1; -
DR      EMBL; AE016756; AAN79032.1; -
DR      EMBL; AE005223; AAG54788.1; -
DR      EMBL; AP002551; BAB33915.1; -
DR      EMBL; AE015070; AAN42039.1; -
DR      PIR; A48709; A48709.
DR      PIR; D90690; D90690.
DR      PIR; H85540; H85540.
DR      HGSP; P32168; I894.
DR      EcoGene; EGI0159; clpX.
DR      HAMAP; MF 00175; -; 1.
DR      InterPro; IPR003593; AAA_ATPase.
```

```

DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR004487; ClpX.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00382; clpX; 1.
KW Chaperone; ATP-binding; Zinc-finger; Heat shock; Complete proteome.
FT INIT MET 0 C4-TYPE.
FT ZN_FING 14 39 ATP (POTENTIAL).
FT NP_BIND 118 125 IGFGATV -> HMCWRS (IN REF. 2).
FT CONFLICT 267 273 7FGDS971FFDCDF9A CRC64;
SQ SEQUENCE 423 AA; 46224 MW; 7FGDS971FFDCDF9A CRC64;

Query Match 1.3%; Score 7; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VYIDEID 90
DB 180 VYIDEID 186

RESULT 53
CLPX_SALTY
ID CLPX_SALTY STANDARD; PRT; 423 AA.
AC Q8Z8V1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE ATP-dependent Clp protease ATP-binding subunit clpX.
GN CLPX OR STY0491 OR T2411.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davis R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Hogue A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RA "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RA "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC It directs the protease to specific substrates. Can perform
CC chaperone functions in the absence of clpP (By similarity).
CC -!- SUBUNIT: Heterodimer of clpP and clpX (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL627266; CAD08908.1; -

```

```

DR EMBL; AE016842; AAO70001.1; -
DR HAMAP; MF 00175; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR004487; ClpX.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00382; clpX; 1.
KW Chaperone; ATP-binding; Zinc-finger; Complete proteome.
FT ZN_FING 15 40 C4-TYPE.
FT NP_BIND 119 126 ATP (POTENTIAL).
FT CONFLICT 119 126 ATP (POTENTIAL).
SQ SEQUENCE 423 AA; 46148 MW; 27D183B952C65454 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VYIDEID 90
DB 181 VYIDEID 187

RESULT 54
CLPX_SALTY
ID CLPX_SALTY STANDARD; PRT; 423 AA.
AC Q8Z8C0; O3LC06;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE ATP-dependent Clp protease ATP-binding subunit clpX.
GN CLPX OR STM0449.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=x3306;
RX MEDLINE=21189233; PubMed=11292737;
RA Yamamoto T., Sashinami H., Takaya A., Tomoyasu T., Matsui H.,
RA Kikuchi Y., Hanawa T., Kamiya S., Nakane A.;
RA "Disruption of the genes for ClpXP protease in Salmonella enterica
serovar Typhimurium results in persistent infection in mice, and
development of persistence requires endogenous gamma interferon and
tumor necrosis factor alpha.";
RL Infect. Immun. 69:3164-3174(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC It directs the protease to specific substrates. Can perform
CC chaperone functions in the absence of clpP (By similarity).
CC -!- SUBUNIT: Heterodimer of clpP and clpX (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB033628; BAA94669.1; -

```

DR EMBL; AE008716; AAL19404.1; -;
 DR StvGene; SG????; clpX.
 DR HAMAP; MF_00175; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_cent.
 DR InterPro; IPR004487; ClpX.
 DR Pfam; PF00004; AAA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR00382; clpX; 1.
 KW Chaperone; ATP-binding; Zinc-finger; Complete proteome.
 FT ZN_FING 15 40 C4-TYPE.
 FT NP_BIND 119 126 ATP (POTENTIAL).
 SQ SEQUENCE 423 AA; 46175 MW; DBA2250FC755B2D9 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 423;
 Best Local Similarity 100.0%; Pred.No.62;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 84 VYIDEID 90
 DB 181 VYIDEID 187
 |||||

RESULT 55
 CLPX_YEREN STANDARD; PRT; 423 AA.
 ID CLPX_YEREN
 AC Q33873;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ATP-dependent Clp protease ATP-binding subunit clpX.
 GN CLPX.
 OS Yersinia enterocolitica
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=630;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8081C / Serotype O:8;
 RX MEDLINE=98043541; PubMed=9383193;
 RA Pederson K.J., Carlson S., Pierson D.E.;
 RT "The ClpX protein, a subunit of the Clp protease, modulates ail gene
 RT expression in Yersinia enterocolitica.";
 RL Mol. Microbiol. 26:99-107(1997).
 CC -!- FUNCTION: ATP-DEPENDENT SPECIFICITY COMPONENT OF THE CLP PROTEASE.
 CC IT DIRECTS THE PROTEASE TO SPECIFIC SUBSTRATES. CAN PERFORM
 CC CHAPERONE FUNCTIONS IN THE ABSENCE OF CLPP (BY SIMILARITY).
 CC -!- SUBUNIT: Heterodimer of clpX and clpX (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U66330; AAC45783.1; -;
 DR HSSP; P32168; 1E94.
 DR HAMAP; MF_00175; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_cent.
 DR InterPro; IPR001270; Chaprinin_clpX/B.
 DR InterPro; IPR004487; ClpX.
 DR Pfam; PF00004; AAA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR00382; clpX; 1.
 KW Chaperone; ATP-binding; Zinc-finger.
 FT ZN_FING 15 40 C4-TYPE.
 FT NP_BIND 119 126 ATP (POTENTIAL).
 SQ SEQUENCE 423 AA; 46168 MW; 958D073D22A9959E CRC64;

Query Match 1.3%; Score 7; DB 1; Length 423;
 Best Local Similarity 100.0%; Pred.No.62;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 84 VYIDEID 90
 DB 181 VYIDEID 187
 |||||

RESULT 56
 CLPX_YERPE STANDARD; PRT; 423 AA.
 ID CLPX_YERPE
 AC Q8ZC66;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ATP-dependent Clp protease ATP-binding subunit clpX.
 GN CLPX OR YPO3156 OR Y1028.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KIM5 / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of Yersinia pestis KIM.";
 RL J. Bacteriol. 184:4601-4611(2002).
 CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
 CC It directs the protease to specific substrates. Can perform
 CC chaperone functions in the absence of clpP (By similarity).
 CC -!- SUBUNIT: Heterodimer of clpP and clpX (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ414155; CAC92391.1; -;
 DR EMBL; AE013706; AAM84609.1; -;
 DR PIR; AD0383; AD0383.
 DR HAMAP; MF_00175; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_cent.
 DR InterPro; IPR004487; ClpX.
 DR Pfam; PF00004; AAA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR00382; clpX; 1.
 KW Chaperone; ATP-binding; Zinc-finger; Complete proteome.
 FT ZN_FING 15 40 C4-TYPE.
 FT NP_BIND 119 126 ATP (POTENTIAL).
 FT NP_BIND 119 126 ATP (POTENTIAL).

```

SQ SEQUENCE 423 AA; 46032 MW; 3D31540123D70D52 CRC64;
Query Match 1.3%; Score 7; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VYIDEID 90
DB 181 VYIDEID 187

RESULT 57
CLPX_RALSO
ID CLPX_RALSO STANDARD; PRT; 424 AA.
AC Q8XYP6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE ATP-dependent Clp protease ATP-binding subunit clpX.
GN CLPX OR RSC1712 OR R502902.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choiane N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguler P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC It directs the protease to specific substrates. Can perform
CC chaperone functions in the absence of clpP (By similarity).
CC -!- SUBUNIT: Heterodimer of clpP and clpX (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL646066; CAD15414.1; -.
CC HAMAP; MF_00175; -.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003959; AAA_ATPase_cent.
CC InterPro; IPR004487; ClpX.
CC Pfam; PF00004; AAA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR00382; clpX; 1.
CC Chaperone; ATP-binding; Zinc-finger; Complete proteome.
CC ZN_FING 15 40 C4-TYPE.
CC FT NP_BIND 122 129 ATP (POTENTIAL).
CC SEQUENCE 424 AA; 46403 MW; 1965182ED59BEAC0 CRC64;

DR EMBL; AL646066; CAD15414.1; -.
DR HAMAP; MF_00175; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR004487; ClpX.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00382; clpX; 1.
DR Chaperone; ATP-binding; Zinc-finger; Complete proteome.
DR ZN_FING 15 40 C4-TYPE.
DR FT NP_BIND 122 129 ATP (POTENTIAL).
DR SEQUENCE 424 AA; 46403 MW; 1965182ED59BEAC0 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 424;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VYIDEID 90
DB 184 VYIDEID 190

RESULT 58
CLPX_PSEAE
ID CLPX_PSEAE STANDARD; PRT; 426 AA.
AC Q912U0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE ATP-dependent Clp protease ATP-binding subunit clpX.
GN CLPX OR PA1802.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;

```



```

RN  SEQUENCE FROM N.A.
RP  STRAIN=ATCC 15692 / PAOI;
RX  MEDLINE=20437337; PubMed=10984043;
RA  Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA  Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA  Garber R.L., Golltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA  Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA  Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA  Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V., an
RT  "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT  opportunistic pathogen.";
RL  Nature 406:959-964 (2000).
CC  -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC  It directs the protease to specific substrates. Can perform
CC  chaperone functions in the absence of clpp (By similarity).
CC  -!- SUBUNIT: Heterodimer of clpp and clpx (By similarity).
CC  -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE004606; AAC05191.1; -.
DR  PIR; F83420; F83420.
DR  HSP; P32168; IE94.
DR  HAMAP; MF 00175; -.
DR  InterPro; IPR003593; AAA_ATPase.
DR  InterPro; IPR003959; AAA_ATPase_cent.
DR  InterPro; IPR004487; ClpX.
DR  Pfam; PF00004; AAA; 1.
DR  SMART; SM00382; AAA; 1.
DR  TIGRFAMs; TIGR00382; clpx; 1.
KW  Chaperone; ATP-binding; Zinc-finger; Complete proteome.
FT  ZN_FING 16 41 C4-TYPE.
FT  NP_BIND 120 127 ATP (POTENTIAL).
SQ  SEQUENCE 426 AA; 46991 MW; 03E7C51E9962CE84 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 426;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VYIDEID 90
DB 182 VYIDEID 188

RESULT 60
CLPX_VIBCH STANDARD; PRT; 426 AA.
AC Q9KQ57;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE ATP-dependent Clp protease (Last annotation update)
DE ATP-dependent Clp protease ATP-binding subunit clpx.
GN CLPX OR VC1921.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bacs S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,

```

```

RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483 (2000).
CC  -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC  It directs the protease to specific substrates. Can perform
CC  chaperone functions in the absence of clpp (By similarity).
CC  -!- SUBUNIT: Heterodimer of clpp and clpx (By similarity).
CC  -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE004268; AAF95069.1; -.
DR  PIR; F82139; F82139.
DR  HSP; P32168; IE94.
DR  TIGR; VC1921; -.
DR  HAMAP; MF 00175; -.
DR  InterPro; IPR003593; AAA_ATPase.
DR  InterPro; IPR003959; AAA_ATPase_cent.
DR  InterPro; IPR004487; ClpX.
DR  Pfam; PF00004; AAA; 1.
DR  SMART; SM00382; AAA; 1.
DR  TIGRFAMs; TIGR00382; clpx; 1.
KW  Chaperone; ATP-binding; Zinc-finger; Complete proteome.
FT  ZN_FING 16 41 C4-TYPE.
FT  NP_BIND 121 128 ATP (POTENTIAL).
SQ  SEQUENCE 426 AA; 46490 MW; 9EA336B942EFA9D4 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 426;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VYIDEID 90
DB 183 VYIDEID 189

RESULT 61
CLPX_VIBPA STANDARD; PRT; 426 AA.
AC Q87R79;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE ATP-dependent Clp protease ATP-binding subunit clpx.
GN CLPX OR VP0918.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749 (2003).
CC  -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC  It directs the protease to specific substrates. Can perform
CC  chaperone functions in the absence of clpp (By similarity).
CC  -!- SUBUNIT: Heterodimer of clpp and clpx (By similarity).
CC  -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC  -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AP005076; BAC59181.1; -;
 DR HAMAP; MF_00175; -; 1.
 KW Chapterone; ATP-binding; Zinc-finger; Complete proteome.
 FT ZN_FING 16 41 C4-TYPE.
 FT NP_BIND 121 128 ATP (POTENTIAL).
 SQ SEQUENCE 426 AA; 46693 MW; A4E899E992DD1668 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 426;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VYIDEID 90
 DB 183 VYIDEID 189

RESULT 62

CLPX_VIBVU STANDARD; PRT; 426 AA.
 ID CLPX_VIBVU
 AC Q8DG27;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE ATP-dependent Clp protease ATP-binding subunit clpX.
 GN CLPX OR VJ10022.
 OS Vibrio vulnificus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=672;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CMCP6;
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
 RA Choy H.E.;
 RT "Complete genome sequence of *Vibrio vulnificus* CMCP6";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
 CC It directs the protease to specific substrates. Can perform
 CC chaperone functions in the absence of clpP (By similarity).
 CC -!- SUBUNIT: Heterodimer of clpP and clpX (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE016797; AAC08566.1; -;
 DR HAMAP; MF_00175; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_central.
 DR Pfam; PF00004; AAA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRfams; TIGR00382; clpX; 1.

KW Chapterone; ATP-binding; Zinc-finger; Complete proteome.
 FT ZN_FING 16 41 C4-TYPE.
 FT NP_BIND 121 128 ATP (POTENTIAL).
 SQ SEQUENCE 426 AA; 46542 MW; D17D84DFAC37CA98 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 426;
 Best Local Similarity 100.0%; Pred. No. 62;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 84 VYIDEID 90
 DB 183 VYIDEID 189

RESULT 63

CLPX_XYLFA STANDARD; PRT; 426 AA.
 ID CLPX_XYLFA
 AC Q9PE40;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ATP-dependent Clp protease ATP-binding subunit clpX.
 GN CLPX OR Xfl188.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=2371;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=9a5c;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carver H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Monck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*";
 RL Nature 406:151-159(2000).
 CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
 CC It directs the protease to specific substrates. Can perform
 CC chaperone functions in the absence of clpP (By similarity).
 CC -!- SUBUNIT: Heterodimer of clpP and clpX (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE003953; AAF83998.1; -;
 DR PIR; B82712; B82712.
 DR HSP; P32168; IE94.
 DR HAMAP; MF_00175; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_central.
 DR InterPro; IPR004487; ClpX.

```

DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00382; clpX; 1.
KW Chaperone; ATP-binding; Zinc-finger; Complete proteome.
FT ZN_FING 18 43 C4-TYPE.
FT NP_BIND 121 128 ATP (POTENTIAL).
SQ SEQUENCE 426 AA; 47076 MW; C64D856C0D666BAE CRC64;

Query Match 1.3%; Score 7; DB 1; Length 426;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 VYIDEID 90
Db 183 VYIDEID 189

RESULT 64
CLPX_BUCAP STANDARD; PRT; 427 AA.
AC Q8K989;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE ATP-dependent Clp protease ATP-binding subunit clpX.
GN CLPX OR BUSG460.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tamas L., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;
RT "50 million years of genomic stasis in endosymbiotic bacteria.";
RL Science 296:2376-2379 (2002).
CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC It directs the protease to specific substrates. Can perform
CC chaperone functions in the absence of clpP (By similarity).
CC -!- SUBUNIT: Heterodimer of clpP and clpX (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
-----
DR EMBL; AB014122; AM68003.1; -.
DR HAMAP; MF_00175; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00382; clpX; 1.
KW Chaperone; ATP-binding; Zinc-finger.
FT ZN_FING 15 40 C4-TYPE.
FT NP_BIND 121 128 ATP (POTENTIAL).
SQ SEQUENCE 427 AA; 47930 MW; 14BA915C06F27598 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 VYIDEID 90
Db 183 VYIDEID 189

RESULT 65
CLPX_MYXXA STANDARD; PRT; 427 AA.
AC Q9X5N1;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE ATP-dependent Clp protease ATP-binding subunit clpX.
GN CLPX.
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DZEL;
RA Ueki T., Inouye S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ATP-DEPENDENT SPECIFICITY COMPONENT OF THE CLP PROTEASE.
CC IT DIRECTS THE PROTEASE TO SPECIFIC SUBSTRATES. CAN PERFORM
CC CHAPERONE FUNCTIONS IN THE ABSENCE OF CLPP (BY SIMILARITY).
CC -!- SUBUNIT: Heterodimer of clpP and clpX (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
-----
DR EMBL; AF127082; AAD31003.1; -.
DR HAMAP; P32168; IE94.
DR HAMAP; MF_00175; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00382; clpX; 1.
KW Chaperone; ATP-binding; Zinc-finger.
FT ZN_FING 16 41 C4-TYPE.
FT NP_BIND 120 127 ATP (POTENTIAL).
SQ SEQUENCE 427 AA; 47156 MW; 2B49D41F659F5F8C CRC64;

Query Match 1.3%; Score 7; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 VYIDEID 90
Db 182 VYIDEID 188

RESULT 66
CLPX_XANAC STANDARD; PRT; 428 AA.
AC Q8PNI4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE ATP-dependent Clp protease ATP-binding subunit clpX.
GN CLPX OR XAC1079.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,

```

```

RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.F.A., L.P.,
RA Camarotte G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC It directs the protease to specific substrates. Can perform
CC chaperone functions in the absence of clpp (By similarity).
CC -!- SUBUNIT: Heterodimer of clpp and clpx (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE011737; AM035957.1; -.
DR HAMAP; MF_00175; -.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR004487; ClpX.
DR Pfam; PF00004; AAA; 1.
DR TIGRFAMs; TIGR00382; clpx; 1.
KW Chaperone; ATP-binding; Zinc-finger; Complete proteome.
FT ZN FING 18 43 C4-TYPE.
FT NP BIND 121 128 ATP (POTENTIAL).
SQ SEQUENCE 428 AA; 47135 MW; 792394511P5C188E CRC64;

Query Match 1.3%; Score 7; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 VYIDEID 90
Db 183 VYIDEID 189

RESULT 67
CLPX_XANCP
ID CLPX_XANCP STANDARD; PRT; 428 AA.
AC Q8PY5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP-dependent Clp protease ATP-binding subunit clpx.
GN CLPX OR XCC0976
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.F.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,

```

```

RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC It directs the protease to specific substrates. Can perform
CC chaperone functions in the absence of clpp (By similarity).
CC -!- SUBUNIT: Heterodimer of clpp and clpx (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE012195; AAM40281.1; -.
DR HAMAP; MF_00175; -.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR004487; ClpX.
DR Pfam; PF00004; AAA; 1.
DR TIGRFAMs; TIGR00382; clpx; 1.
KW Chaperone; ATP-binding; Zinc-finger; Complete proteome.
FT ZN FING 18 43 C4-TYPE.
FT NP BIND 121 128 ATP (POTENTIAL).
SQ SEQUENCE 428 AA; 47136 MW; AFD064703FD7F73F CRC64;

Query Match 1.3%; Score 7; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 VYIDEID 90
Db 183 VYIDEID 189

RESULT 68
CLPX_BUCAI
ID CLPX_BUCAI STANDARD; PRT; 429 AA.
AC P57548;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP-dependent Clp protease ATP-binding subunit clpx.
GN CLPX OR BU476.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS."
RL Nature 407:81-86(2000).
CC -!- FUNCTION: ATP-dependent specificity component of the Clp protease.
CC It directs the protease to specific substrates. Can perform
CC chaperone functions in the absence of clpp (By similarity).
CC -!- SUBUNIT: Heterodimer of clpp and clpx (By similarity).

```

CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001119; BAB13173.1; -;
DR HSSP; P32168; IE94.
DR HAMAP; MF_00175; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR004487; ClpX.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRfam; TIGR00382; clpX; 1.
KW Chaperone; ATP-binding; Zinc-finger; Complete proteome.
FT ZN_FING 15 40 C4-TYPE
FT NP_BIND 124 131 ATP (POTENTIAL).
SQ SEQUENCE 429 AA; 48009 MW; A859C79A8230509F CRC64;

Query Match 1.3%; Score 7; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 VYIDRID 90
Db 186 VYIDRID 192
|||||

RESULT 69
DHAL_PSSP
ID DHAL_PSSP STANDARD; PRT; 488 AA.
AC P33008;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable aldehyde dehydrogenase (EC 1.2.1.3).
GN TERPE.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92332528; PubMed=1629218;
RA Peterson J.A., Lu J.-Y., Geiselsoder J., Graham-Lorence S.,
RA Carmona C., Witney P., Lorence M.C.;
RT "Cytochrome P-450terp. Isolation and purification of the protein and
RT cloning and sequencing of its operon."
RL J. Biol. Chem. 267:14193-14203(1992).
CC -!- FUNCTION: INVOLVED IN AN ALPHA-TERPENEOL OXIDATION SYSTEM.
CC -!- CATALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2)O = an acid + NADH.
CC -!- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M91440; AAA25995.1; -;
DR PIR; C42971; S27652.
DR HSSP; P51977; 1BXS.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Oxidoreductase; NAD.

FT NP_BIND 240 245 NAD (ADP PART) (BY SIMILARITY).
FT ACT_SITE 262 262 BY SIMILARITY.
FT ACT_SITE 296 296 BY SIMILARITY.
SQ SEQUENCE 488 AA; 52156 MW; 2144AE1EB173C854 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 488;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 468 HVEQARA 474
Db 61 HVEQARA 67
|||||

RESULT 70
C6B1_PAPPO
ID C6B1_PAPPO STANDARD; PRT; 498 AA.
AC Q04552; Q04553; Q27878;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 6B1 (EC 1.14.14.1) (CYP6B1) (CYP6B1V1/CYP6B1V2/
DE CYP6B1V3).
GN CYP6B1.
OS Papilio polyxenes (black swallowtail butterfly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Papilionoidea; Papilionidae; Papilioninae; Papilio.
OX NCBI_TaxID=7146;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-25 AND 401-406.
RC TISSUE=Midgut;
RX MEDLINE=93066355; PubMed=1279697;
RA Cohen M.B., Schuler M.A., Berenbaum M.R.;
RT "A host-inducible cytochrome P-450 from a host-specific caterpillar:
RT molecular cloning and evolution."
RL Proc. Natl. Acad. Sci. U.S.A. 89:10920-10924(1992).
RN [2]
RP SEQUENCE FROM N.A. (CYP6B1V3).
RC TISSUE=Midgut;
RX MEDLINE=94344788; PubMed=8065937;
RA Prapaipong H.H., Berenbaum M.M., Schuler M.M.;
RT "Transcriptional regulation of the Papilio polyxenes CYP6B1 gene."
RL Nucleic Acids Res. 22:3210-3217(1994).
CC -!- FUNCTION: ENABLES THE INSECT TO FEED ON FURANOCUMARIN-PRODUCING
CC PLANTS AND EVOLVED AS AN ADAPTATION FOR DETOXIFICATION OF
CC XANTHOTOXIN AND OTHER FURANOCUMARINS.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- TISSUE SPECIFICITY: MIDGUT MICROSOME.
CC -!- INDUCTION: BY XANTHOTOXIN, A SECONDARY METABOLITE ABUNDANT IN THE
CC HOST PLANTS OF THIS SPECIALIZED HERBIVORE.
CC -!- POLYMORPHISM: THE SEQUENCE SHOWN IS THAT OF 6B1-1, 6B1-2 SEEMS
CC TO DIFFER IN 9 POSITIONS AND IS PROBABLY AN ALLELE.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M80828; AAA29789.1; -;
DR PIR; M83117; AAA29790.1; -;
DR EMBL; Z29624; CAA82732.1; -;
DR EMBL; U05037; AAA16154.1; -;
DR PIR; S48058; S48058.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.

DR PROSITE, PS00086; CYTOCHROME P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum; Polymorphism.
 FT METAL 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT VARIANT 24 N -> D (IN 6B1-2 AND 6B1-3).
 FT VARIANT 155 NS -> KC (IN 6B1-2 AND 6B1-3).
 FT VARIANT 243 K -> R (IN 6B1-2).
 FT VARIANT 285 A -> S (IN 6B1-2).
 FT VARIANT 293 I -> V (IN 6B1-2).
 FT VARIANT 458 M -> V (IN 6B1-2).
 FT VARIANT 475 P -> E (IN 6B1-2).
 FT VARIANT 495 L -> I (IN 6B1-2).
 SQ SEQUENCE 498 AA; 57483 MW; 043A849CA0990153 CRC64;

Query Match 1.3%; Score 7; DB 1; Length 498;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 461 VLEMRKG 467
 DB 245 VLEMRKG 251
 |||||

RESULT 71
 GNTK_BACSU STANDARD; PRT; 513 AA.
 AC P12011;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glucuronokinase (EC 2.7.1.12) (Glucuronate kinase).
 GN GNTK.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87008613; PubMed=3020045;
 RA Fujita Y., Fujita T., Miwa Y., Nihashi J., Aratani Y.;
 RT "Organization and transcription of the glucuronate operon, gnt, of
 RT Bacillus subtilis.";
 RL J. Biol. Chem. 261:13744-13753(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / BGSC1A1;
 RX MEDLINE=96093926; PubMed=7584049;
 RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
 RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis
 RT genome between the gnt and iol operons.";
 RL DNA Res. 2:61-69(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koeter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Maeda S., Maue C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Neone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + D-glucuronate = ADP + 6-phospho-D-
 CC Glucuronate.
 CC -1- ENZYME REGULATION: CATABOLITE REPRESSION BY GLUCONATE.
 CC -1- PATHWAY: Glucuronate utilization.
 CC -1- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCONOKINASE /
 CC GLYCEROKINASE / XYLULOSE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J02584; AA56925.1; -;
 DR EMBL; AB005554; BAA21578.1; -;
 DR EMBL; Z99124; CAB16043.1; -;
 DR PIR; B26190; B26190.
 DR Subtilist; BGI0649; gntK.
 DR InterPro; IPR000577; FGGY_kin.
 DR InterPro; IPR006002; Glucuronate_kinase.
 DR Pfam; PF02782; FGGY; 1.
 DR TIGRFAMs; TIGR01314; gntK_FGGY; 1.
 DR PROSITE; PS00445; FGGY_KINASES 2; 1.
 DR PROSITE; PS00933; FGGY_KINASES 1; 1.
 KW Glucuronate utilization; Transferase; Kinase; Complete proteome.
 SQ SEQUENCE 513 AA; 57169 MW; 351D44064C7AC3ED CRC64;

Query Match 1.3%; Score 7; DB 1; Length 513;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 FINLSRL 202
 DB 488 FINLSRL 494
 |||||

RESULT 72
 HEMA_SV5 STANDARD; PRT; 565 AA.
 AC P04850;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemagglutinin-neuraminidase (EC 3.2.1.18).
 GN HN.
 OS Simian virus 5 (strain W3) (SV5).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
 OX NCBI_TaxID=11208;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85135055; PubMed=3973974;
 RA Hiebert S.W., Paterson R.G., Lamb R.A.;
 RT "Hemagglutinin-neuraminidase protein of the paramyxovirus simian
 RT virus 5: nucleotide sequence of the mRNA predicts an N-terminal
 RT membrane anchor.";
 RL J. Virol. 54:1-6(1985).
 RN [2]

SEQUENCE FROM N.A.
 MEDLINE=92113554; PubMed=1765772;
 Baty D.U., Southern J.A., Randall R.E.;
 "Sequence comparison between the haemagglutinin-neuraminidase genes
 of simian, canine and human isolates of simian virus 5.";
 J. Gen. Virol. 72:3103-3107(1991).
 -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS
 TO CELL RECEPTORS AND FOR INITIATING INFECTION. NEURAMINIDASE
 ACTIVITY HELPS THE EFFICIENT SPREAD OF THE VIRUS BY DISSOCIATING
 THE MATURE VIRIONS FROM THE NEURAMINIC ACID CONTAINING
 GLYCOPROTEINS.
 -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
 alpha-(2->8)-glycosidic linkages of terminal sialic residues in
 oligosaccharides, glycoproteins, glycolipids, colominic acid and
 synthetic substrates.
 -1- SUBCELLULAR LOCATION: EXTERNAL, ANCHORED TO THE ENVELOPE BY ITS
 N-TERMINAL HYDROPHOBIC SEQUENCE.
 -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES HEMAGGLUTININ-
 NEURAMINIDASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; S76876; AAB21114.1; -;
 EMBL; K02870; AAA47878.1; -;
 EMBL; AF052755; AAC95517.1; -;
 PIR; A00879; HNNZSV.
 InterPro; IPR000665; Hem-neuramndase.
 Pfam; PF00423; HN; 1.
 Hydrolase; Hemagglutinin; Envelope protein; Glycoprotein;
 Transmembrane.
 DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 21 40 POTENTIAL.
 DOMAIN 41 565 EXTRACELLULAR (POTENTIAL).
 CARBOHYD 110 110 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 139 139 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
 SEQUENCE 565 AA; 62204 MW; 30A4A7068A25B563 CRC64;
 Query Match 1.3%; Score 7; DB 1; Length 565;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 108 QVPRPGT 114
 DB 439 QVPRPGT 445
 RESULT 73
 HEMA_SV5CM STANDARD; PRT; 565 AA.
 AC P28883;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemagglutinin-neuraminidase (EC 3.2.1.18).
 GN HN.
 OS Simian virus 5 (isolate canine/CPI-) (SV5).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
 OX NCBI_TaxID=31609;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92113554; PubMed=1765772;
 Baty D.U., Southern J.A., Randall R.E.;
 "Sequence comparison between the haemagglutinin-neuraminidase genes
 of simian, canine and human isolates of simian virus 5.";
 J. Gen. Virol. 72:3103-3107(1991).
 -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS
 TO CELL RECEPTORS AND FOR INITIATING INFECTION. NEURAMINIDASE
 ACTIVITY HELPS THE EFFICIENT SPREAD OF THE VIRUS BY DISSOCIATING
 THE MATURE VIRIONS FROM THE NEURAMINIC ACID CONTAINING
 GLYCOPROTEINS.
 -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
 alpha-(2->8)-glycosidic linkages of terminal sialic residues in
 oligosaccharides, glycoproteins, glycolipids, colominic acid and
 synthetic substrates.
 -1- SUBCELLULAR LOCATION: EXTERNAL, ANCHORED TO THE ENVELOPE BY ITS
 N-TERMINAL HYDROPHOBIC SEQUENCE.
 -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES HEMAGGLUTININ-
 NEURAMINIDASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; S76876; AAB21114.1; -;
 EMBL; K02870; AAA47878.1; -;
 EMBL; AF052755; AAC95517.1; -;
 PIR; A00879; HNNZSV.
 InterPro; IPR000665; Hem-neuramndase.
 Pfam; PF00423; HN; 1.
 Hydrolase; Hemagglutinin; Envelope protein; Glycoprotein;
 Transmembrane.
 DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 21 40 POTENTIAL.
 DOMAIN 41 565 EXTRACELLULAR (POTENTIAL).
 CARBOHYD 110 110 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 139 139 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
 SEQUENCE 565 AA; 62204 MW; 30A4A7068A25B563 CRC64;
 Query Match 1.3%; Score 7; DB 1; Length 565;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 108 QVPRPGT 114
 DB 439 QVPRPGT 445
 RESULT 74
 HEMA_SV5CP STANDARD; PRT; 565 AA.
 AC P28884;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemagglutinin-neuraminidase (EC 3.2.1.18).
 GN HN.
 OS Simian virus 5 (isolate canine/CPI+) (SV5).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
 OX NCBI_TaxID=31608;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92113554; PubMed=1765772;
 Baty D.U., Southern J.A., Randall R.E.;
 "Sequence comparison between the haemagglutinin-neuraminidase genes
 of simian, canine and human isolates of simian virus 5.";
 J. Gen. Virol. 72:3103-3107(1991).
 -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS
 TO CELL RECEPTORS AND FOR INITIATING INFECTION. NEURAMINIDASE
 ACTIVITY HELPS THE EFFICIENT SPREAD OF THE VIRUS BY DISSOCIATING
 THE MATURE VIRIONS FROM THE NEURAMINIC ACID CONTAINING
 GLYCOPROTEINS.
 -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
 alpha-(2->8)-glycosidic linkages of terminal sialic residues in
 oligosaccharides, glycoproteins, glycolipids, colominic acid and
 synthetic substrates.
 -1- SUBCELLULAR LOCATION: EXTERNAL, ANCHORED TO THE ENVELOPE BY ITS
 N-TERMINAL HYDROPHOBIC SEQUENCE.
 -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES HEMAGGLUTININ-
 NEURAMINIDASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; S76876; AAB21114.1; -;
 EMBL; K02870; AAA47878.1; -;
 EMBL; AF052755; AAC95517.1; -;
 PIR; A00879; HNNZSV.
 InterPro; IPR000665; Hem-neuramndase.
 Pfam; PF00423; HN; 1.
 Hydrolase; Hemagglutinin; Envelope protein; Glycoprotein;
 Transmembrane.
 DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 21 40 POTENTIAL.
 DOMAIN 41 565 EXTRACELLULAR (POTENTIAL).
 CARBOHYD 110 110 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 139 139 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
 SEQUENCE 565 AA; 62279 MW; 4DA94276001FBD3C CRC64;
 Query Match 1.3%; Score 7; DB 1; Length 565;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 108 QVPRPGT 114
 DB 439 QVPRPGT 445

J. Gen. Virol. 72:3103-3107(1991).
 -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS
 TO CELL RECEPTORS AND FOR INITIATING INFECTION. NEURAMINIDASE
 ACTIVITY HELPS THE EFFICIENT SPREAD OF THE VIRUS BY DISSOCIATING
 THE MATURE VIRIONS FROM THE NEURAMINIC ACID CONTAINING
 GLYCOPROTEINS.
 -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
 alpha-(2->8)-glycosidic linkages of terminal sialic residues in
 oligosaccharides, glycoproteins, glycolipids, colominic acid and
 synthetic substrates.
 -1- SUBCELLULAR LOCATION: EXTERNAL, ANCHORED TO THE ENVELOPE BY ITS
 N-TERMINAL HYDROPHOBIC SEQUENCE.
 -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES HEMAGGLUTININ-
 NEURAMINIDASE FAMILY.
 PIR; J01305; HNNZC1.
 InterPro; IPR000665; Hem-neuramndase.
 Pfam; PF00423; HN; 1.
 Hydrolase; Hemagglutinin; Envelope protein; Glycoprotein;
 Transmembrane.
 DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 21 40 POTENTIAL.
 DOMAIN 41 565 EXTRACELLULAR (POTENTIAL).
 CARBOHYD 110 110 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 139 139 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
 SEQUENCE 565 AA; 62279 MW; 4DA94276001FBD3C CRC64;
 Query Match 1.3%; Score 7; DB 1; Length 565;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 108 QVPRPGT 114
 DB 439 QVPRPGT 445
 RESULT 74
 HEMA_SV5CP STANDARD; PRT; 565 AA.
 AC P28884;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemagglutinin-neuraminidase (EC 3.2.1.18).
 GN HN.
 OS Simian virus 5 (isolate canine/CPI+) (SV5).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
 OX NCBI_TaxID=31608;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92113554; PubMed=1765772;
 Baty D.U., Southern J.A., Randall R.E.;
 "Sequence comparison between the haemagglutinin-neuraminidase genes
 of simian, canine and human isolates of simian virus 5.";
 J. Gen. Virol. 72:3103-3107(1991).
 -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS
 TO CELL RECEPTORS AND FOR INITIATING INFECTION. NEURAMINIDASE
 ACTIVITY HELPS THE EFFICIENT SPREAD OF THE VIRUS BY DISSOCIATING
 THE MATURE VIRIONS FROM THE NEURAMINIC ACID CONTAINING
 GLYCOPROTEINS.
 -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
 alpha-(2->8)-glycosidic linkages of terminal sialic residues in
 oligosaccharides, glycoproteins, glycolipids, colominic acid and
 synthetic substrates.
 -1- SUBCELLULAR LOCATION: EXTERNAL, ANCHORED TO THE ENVELOPE BY ITS
 N-TERMINAL HYDROPHOBIC SEQUENCE.
 -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES HEMAGGLUTININ-
 NEURAMINIDASE FAMILY.
 PIR; J01306; HNNZC2.
 InterPro; IPR000665; Hem-neuramndase.
 Pfam; PF00423; HN; 1.

KW Hydrolase; Hemagglutinin; Envelope protein; Glycoprotein; Transmembrane.
FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 21 40 POTENTIAL.
FT DOMAIN 41 565 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 565 AA; 62307 MW; 4DA94276AC6F523C CRC64;

Query Match 1.3%; Score 7; DB 1; Length 565;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 QVPRPGT 114
|||||

Db 439 QVPRPGT 445

RESULT 75
HEMA_SV5LN
ID HEMA_SV5LN STANDARD; PRT; 565 AA.
AC P28885;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemagglutinin-neuraminidase (EC 3.2.1.18).
GN HN.
OS Simian virus 5 (isolate human/LN) (SV5).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
OX NCBI_TaxID=31610;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113554; PubMed=1765772;
RA Baty D.O., Southern J.A., Randall R.E.;
RT "Sequence comparison between the haemagglutinin-neuraminidase genes
of simian, canine and human isolates of simian virus 5";
RL J. Gen. Virol. 72:3103-3107(1991).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS
TO CELL RECEPTORS AND FOR INITIATING INFECTION. NEURAMINIDASE
ACTIVITY HELPS THE EFFICIENT SPREAD OF THE VIRUS BY DISSOCIATING
THE MATURE VIRIONS FROM THE NEURAMINIC ACID CONTAINING
GLYCOPROTEINS.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
alpha-(2->8)-glycosidic linkages of terminal sialic residues in
oligosaccharides, glycoproteins, glycolipids, colominic acid and
synthetic substrates.
CC -!- SUBCELLULAR LOCATION: EXTERNAL, ANCHORED TO THE ENVELOPE BY ITS
N-TERMINAL HYDROPHOBIC SEQUENCE.
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES HEMAGGLUTININ-
NEURAMINIDASE FAMILY.
DR PIR; JQ1307; HNNZC3.
DR InterPro; IPR000665; Hem-neuramidase.
DR Pfam; PF00423; HN; 1.
KW Hydrolase; Hemagglutinin; Envelope protein; Glycoprotein; Transmembrane.
FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 21 40 POTENTIAL.
FT DOMAIN 41 565 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 565 AA; 62137 MW; 8F531B8704F281D CRC64;

Query Match 1.3%; Score 7; DB 1; Length 565;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 QVPRPGT 114
|||||

Db 439 QVPRPGT 445

Search completed: February 10, 2004, 13:29:36
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 13:26:36 ; Search time 47 Seconds
(without alignments)
1813.536 Million cell updates/sec

Title: US-09-821-812-5
Perfect score: 537
Sequence: 1 GGLTTTIGRLGVRRLS.....PDHVDTHLIKLRQHPFAML 537

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : A Geneseq 19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	187	34.8	260	22	AA41765 Human polypeptide
2	166	30.9	220	21	AA53386 Human colon cancer
3	141	26.3	142	22	ABG18793 Novel human diagno
4	92	17.1	550	22	ABG18795 Novel human diagno
5	64	11.9	207	22	ABG18794 Novel human diagno
6	38	7.1	90	22	AA33979 Human polypeptide
7	16	3.0	101	22	AAU32896 Novel human secret
8	8	1.5	123	20	AAW93405 Human HEV ORF 3 pr
9	8	1.5	283	23	ABB04717 Human PF1030 prote

10	8	1.5	487	21	AA10901	S. xyloosus DltA pr
11	7	1.3	29	17	AAW00292	Penicillin V amido
12	7	1.3	36	20	AAW97535	Antigenic site of
13	7	1.3	59	23	ABB49653	Listeria monocytog
14	7	1.3	66	21	AB333405	Pinus radiata tran
15	7	1.3	73	23	ABF04937	Human ORFX protein
16	7	1.3	75	22	AAU34866	E. coli cellular p
17	7	1.3	75	22	AAU35457	Haemophilus influe
18	7	1.3	75	22	AAU38266	Salmonella typhi c
19	7	1.3	75	22	AB116053	Human nervous syst
20	7	1.3	82	22	AAU42302	Propionibacterium
21	7	1.3	85	23	ABF35330	Human ORF4303 prot
22	7	1.3	103	20	AAV19803	B. burgdorferi ant
23	7	1.3	133	21	AAV19802	A. burgdorferi ant
24	7	1.3	144	21	AA266620	Arabidopsis thalia
25	7	1.3	144	21	AA486630	Arabidopsis thalia
26	7	1.3	146	21	AA333260	Pinus radiata tran
27	7	1.3	171	22	ABB60190	Drosophila melanog
28	7	1.3	209	22	ABB71753	Drosophila melanog
29	7	1.3	213	21	AA266619	Arabidopsis thalia
30	7	1.3	213	21	AA486629	Arabidopsis thalia
31	7	1.3	215	15	AA560575	House dust mite al
32	7	1.3	215	20	AA255586	D. pteronyssinus a
33	7	1.3	267	19	AAW81726	M. tuberculosis im
34	7	1.3	267	19	AAW64359	Mycobacterium tube
35	7	1.3	267	20	AAV39156	M. tuberculosis an
36	7	1.3	267	20	AAV39013	M. tuberculosis re
37	7	1.3	287	22	ABG07558	Novel human diagno
38	7	1.3	290	21	AA236334	Arabidopsis thalia
39	7	1.3	307	21	AA236333	Arabidopsis thalia
40	7	1.3	309	21	AAV74359	Neisseria meningit
41	7	1.3	309	21	AAV74361	Neisseria meningit
42	7	1.3	310	21	AA236332	Arabidopsis thalia
43	7	1.3	312	21	AAV74360	Neisseria meningit
44	7	1.3	315	21	AAV58367	Anabaena variabil
45	7	1.3	317	23	AB557933	Mouss ischaemic co
46	7	1.3	321	22	ABG20636	Novel human diagno
47	7	1.3	329	23	AB922389	Herbicidally activ
48	7	1.3	330	23	ABP51947	Haemophilus influe
49	7	1.3	340	21	AAV52000	Human ACL1 protein
50	7	1.3	340	21	AAV51629	Human ACL1 protein
51	7	1.3	363	22	AA96616	Putative P. abyss
52	7	1.3	370	21	AA332774	Eucalyptus grandis
53	7	1.3	373	21	AA486637	Arabidopsis thalia
54	7	1.3	373	23	AB847970	Listeria monocytog
55	7	1.3	383	21	AA255383	Arabidopsis thalia
56	7	1.3	384	21	AA486636	Arabidopsis thalia
57	7	1.3	387	22	ABG14658	Novel human diagno
58	7	1.3	392	21	AA486625	Arabidopsis thalia
59	7	1.3	394	21	AA253382	Arabidopsis thalia
60	7	1.3	398	22	AB863456	Drosophila melanog
61	7	1.3	401	22	AB862489	Drosophila melanog
62	7	1.3	403	21	AA486624	Arabidopsis thalia
63	7	1.3	412	21	AA333303	Pinus radiata tran
64	7	1.3	414	24	ABP78893	N. gonorrhoeae ami
65	7	1.3	419	22	AB112386	Human bone marrow
66	7	1.3	420	21	AA486635	Arabidopsis thalia
67	7	1.3	430	21	AA253381	Arabidopsis thalia
68	7	1.3	433	22	AAU48807	Propionibacterium
69	7	1.3	437	19	AAW41733	Arabidopsis chloro
70	7	1.3	439	21	AA486623	Arabidopsis thalia
71	7	1.3	470	24	ABP80005	N. gonorrhoeae ami
72	7	1.3	507	17	AAW00290	Mature Penicillin
73	7	1.3	532	17	AAW00291	Full length Penici
74	7	1.3	544	20	ABY34962	Chlamydia pneumoni
75	7	1.3	565	24	ABJ18511	Human Cryptovirus
76	7	1.3	582	23	ABB94315	Chlamydia pneumoni
77	7	1.3	593	14	AA42479	Mus RYK. Mus mu
78	7	1.3	638	22	AAE00813	Aspergillus niger
79	7	1.3	677	22	ABU53116	Human testee-deriv
80	7	1.3	693	23	ABG70803	Human myocardin 2
81	7	1.3	693	24	ABG74175	Human cardiac-spec
82	7	1.3	770	22	AA310824	Amino acid sequenc

83 Human pendrin SEQ
84 Prostate cancer-as
85 Human Protease and
86 Human paraplegin a
87 S. epidermidis ope
88 Porphyromonas ging
89 Porphyromonas ging
90 Porphyromonas ging
91 Novel human diagno
92 Staphylococcus epi
93 Human polypeptide
94 Human MKL1 protein
95 Tomato pathogen re
96 Tomato pathogen re
97 Human lung-specifi
98 Novel human diagno
99 C. vicina arylphor
100 H. muticus FDR1 pr

ALIGNMENTS

RESULT 1

AA41765
ID AAM41765 standard; Protein; 260 AA.

XX AC AAM41765;
XX DT 22-OCT-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 6696.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0620312.

XX PR 03-AUG-2000; 2000US-0653450.

XX PR 14-SEP-2000; 2000US-0662191.

XX PR 19-OCT-2000; 2000US-0693036.

XX PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX N-PSDB; AAI60921.

XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -

XX Example 2; SEQ ID NO 6696; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC assays for receptor activity, cancer diagnosis and therapy, drug screening,
CC and thrombolytic activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX SQ Sequence 260 AA;

Query Match 34.8%; Score 187; DB 22; Length 260;

Best Local Similarity 100.0%; Pred. No. 5.5e-183;

Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PAAMSEMEPLLLAWSYFRRKFKQCADLCTQMLEKSPYDQAAMILKARALTEMYIDEI 89

DB 17 PAAMSEMEPLLLAWSYFRRKFKQCADLCTQMLEKSPYDQAAMILKARALTEMYIDEI 76

QY 90 DVDOEGTAEMMLDENAJAQPVRPGTSLKLPCTNQTGGPSQAVRPITQAGRPITGFLRPST 149

DB 77 DVDOEGTAEMMLDENAJAQPVRPGTSLKLPCTNQTGGPSQAVRPITQAGRPITGFLRPST 136

QY 150 QSGRPGTMEQAIPTPTAYTARPTITSSSGRVRVLTASMLTSPDGPFINLSRLNLTKYSQ 209

DB 137 QSGRPGTMEQAIPTPTAYTARPTITSSSGRVRVLTASMLTSPDGPFINLSRLNLTKYSQ 196

QY 210 KPILAKA 216

DB 197 KPILAKA 203

RESULT 2

AA53386

ID AAB53386 standard; Protein; 220 AA.

XX AC AAB53386;

XX DT 09-MAR-2001 (first entry)

XX DE Human colon cancer antigen protein sequence SEQ ID NO:926.

XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytostatic; cardioactive; neuroprotective; vulnary;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder.

XX OS Homo sapiens.

XX PN WO200055351-A1.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US05883.

XX PR 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX WPI; 2000-587534/55.

XX N-PSDB; AAC98143.

PT Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PT disorders such as colon cancer -

XX Claim 11; Page 1485; 2104pp; English.

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC human colon cancer antigens can have cytosolic, cardioactive, muscular;
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and
CC vulnary, nephrotropic, antineoplastic and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins
CC may also be used to prevent diseases such as neural disorders, immune
CC system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, wounds, renal disorders, infectious
CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
CC AAB54007 represent sequences used in the exemplification of the present
CC invention.

XX SQ Sequence 220 AA;

Query Match 30.9%; Score 166; DB 21; Length 220;

Best Local Similarity 100.0%; Pred. No. 1.7e-161; Mismatches 0; Indels 0; Gaps 0;

Matches 166; Conservative 0; Indels 0; Gaps 0;

Yy 372 RRLQMGVYNGQLFNNLGLCCFYAQYDMLTSPERALSFAENEEAADVYVNLGHVAVG 431

Db 55 RRLQMGVYNGQLFNNLGLCCFYAQYDMLTSPERALSFAENEEAADVYVNLGHVAVG 114

Yy 432 IGDNLHQCPRALVNNNHAENYNNLAVLEMRKHVEQARALLQTASSLAPHMYEPHF 491

Db 115 IGDNLHQCPRALVNNNHAENYNNLAVLEMRKHVEQARALLQTASSLAPHMYEPHF 174

Yy 492 NFATISDKIGLQRSYVAAQKSEAFDPDVTQHLIKQLRHFAML 537

Db 175 NFATISDKIGLQRSYVAAQKSEAFDPDVTQHLIKQLRHFAML 220

RESULT 3

ABG18793

ID ABG18793 standard; Protein; 142 AA.

XX AC ABG18793;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #18784.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS82980.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID No 49152; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 142 AA;

Query Match 26.3%; Score 141; DB 22; Length 142;

Best Local Similarity 100.0%; Pred. No. 5e-136;

Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yy 46 YFRRRFQLCADLCTQMLEKSPYDQAAWILKARALTEMVYIDEIDVDQEGIAEMMLDENA 105

Db 2 YFRRRFQLCADLCTQMLEKSPYDQAAWILKARALTEMVYIDEIDVDQEGIAEMMLDENA 61

Yy 106 IAQVPRGTSKLPGNTQGTGSPQAVRPITQAGRPITGFLRPSQTSGRPGTMEQAIPTPR 165

Db 62 IAQVPRGTSKLPGNTQGTGSPQAVRPITQAGRPITGFLRPSQTSGRPGTMEQAIPTPR 121

Yy 166 TAYTARPITSSGRFVRLGTA 186

Db 122 TAYTARPITSSGRFVRLGTA 142

RESULT 4

ABG18795

ID ABG18795 standard; Protein; 550 AA.

XX AC ABG18795;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #18786.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX

```
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS82982.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX Claim 20; SEQ ID NO 49154; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 550 AA;
SQ
Query Match 17.1%; Score 92; DB 22; Length 550;
Best Local Similarity 100.0%; Pred. No. 3.4e-85;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 350 ATACGSHFYSDQFEIALRFRVRLQLGQVYNGQLFNNLGLCCFYAQVQYDMLTSLFERAL 409
DB 361 ATACGSHFYSDQFEIALRFRVRLQLGQVYNGQLFNNLGLCCFYAQVQYDMLTSLFERAL 420
QY 410 SLAENEERAAADVWYNLGHVAVGIGDTNLAHQ 441
DB 421 SLAENEERAAADVWYNLGHVAVGIGDTNLAHQ 452
RESULT 5
ABG18794
ID ABG18794 standard; Protein; 207 AA.
XX
AC ABG18794;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #18785.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX
XX 23-AUG-2000; 2000US-0649167.
```

```
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS82981.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX Claim 20; SEQ ID NO 49153; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 207 AA;
SQ
Query Match 11.9%; Score 64; DB 22; Length 207;
Best Local Similarity 100.0%; Pred. No. 7.4e-57;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 70 QAAWILKARALTEMVYIDEIDVDQEGIAEMMLDENATAQVPRPGTSLKLPCTNQTGGPSQ 129
DB 1 QAAWILKARALTEMVYIDEIDVDQEGIAEMMLDENATAQVPRPGTSLKLPCTNQTGGPSQ 60
QY 130 AVRP 133
DB 61 AVRP 64
RESULT 6
AAM39979
ID AAM39979 standard; Protein; 90 AA.
XX
AC AAM39979;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 3124.
XX
KW Human; nototropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
```

PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0682191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AAI59135.
 XX
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PT
 XX Example 4; SEQ ID NO 3124; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 XX Sequence 90 AA;
 SQ
 Query Match 7.1%; Score 38; DB 22; Length 90;
 Best Local Similarity 100.0%; Pred. No. 1.6e-30;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 377 MGIYNGQLFNNLGLCCFYAQOYDWTLSFERALSIAEN 414
 DB 51 MGIYNGQLFNNLGLCCFYAQOYDWTLSFERALSIAEN 88
 RESULT 7
 AAU32896
 ID AAU32896 standard; Protein; 101 AA.
 XX
 XX AAU32896;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #3387.
 XX
 KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX

PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US08656.
 XX
 PR 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-611725/70.
 XX
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 PT
 XX Claim 20; Page 685; 765pp; English.
 XX
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX
 XX Sequence 101 AA;
 SQ
 Query Match 3.0%; Score 16; DB 22; Length 101;
 Best Local Similarity 100.0%; Pred. No. 6.8e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 454 EAYNNLAVLEMRKGHV 469
 DB 30 EAYNNLAVLEMRKGHV 45
 RESULT 8
 AAU93405
 ID AAU93405 standard; Protein; 123 AA.
 XX
 AC AAU93405;
 XX
 DT 11-JUN-1999 (first entry)
 XX
 DE Human HEV ORF 3 protein from strain Hetian.
 XX
 KW Swine hepatitis E virus; HEV; cross-reaction; antibody; human; therapy;
 KW vaccine; immunise; infection; detection; diagnosis; prevention.
 XX
 OS Hepatitis E virus.
 XX
 PN WO9904029-A2.
 XX
 PD 28-JAN-1999.
 XX
 PF 17-JUL-1998; 98WO-US14665.
 XX
 PR 18-JUL-1997; 97US-0053069.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Emerson SU, Meng X, Purcell RH;

XX WPI; 1999-132270/11.
XX
XX New isolated swine hepatitis E virus - used to develop products for
PT the diagnosis, prevention and treatment of hepatitis E virus
PT infection in mammals, particularly humans
XX
XX Example 1; Fig 3B; 70pp; English.
XX
XX This invention describes a swine hepatitis E virus (HEV) and its natural
CC mutants which are capable of cross-reacting with antibodies reactive
CC with a human HEV strain or natural mutants. The HEV and the proteins
CC can be used in vaccines for immunising against HEV infection. The swine
CC HEV can be used in humans to prevent possible infection by human HEV. The
CC swine HEV can also be used as a therapeutic treatment for infection by
CC other strains of HEV. The swine HEV can also be used for the production
CC of antibodies which can be used in therapy, detection and diagnosis. The
CC products can also be used for determining the susceptibility of cells or
CC organs to infection with swine HEV. The swine HEV is particularly useful
CC for the development of agents for the prevention, treatment and detection
CC of human HEV because it is not a human virus and thus can be handled both
CC experimentally and clinically without fear of severe infection and/or
CC contamination.
XX
XX Sequence 123 AA;
SQ
Query Match 1.5%; Score 8; DB 20; Length 123;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 327 MNMSSAA 334
Db 1 MNMSSAA 8
RESULT 9
ID ABB04717 standard; Protein; 283 AA.
XX
XX ABB04717;
AC
XX
XX 11-MAR-2002 (first entry)
DT
XX Human PP1030 protein SEQ ID NO:5.
DE
XX Human; PP1030; cancer suppression.
KW
XX Homo sapiens.
OS
XX CN1313316-A.
PN
XX 19-SEP-2001.
PD
XX 13-MAR-2000; 2000CN-0111990.
PF
XX 13-MAR-2000; 2000CN-0111990.
PR
XX (SHAN-) SHANGHAI INST ONCOLOGY.
PA
XX Gu J, Yang S;
PI
XX WPI; 2002-042194/06.
DR
XX N-PSDB; ABA04452.
DR
XX New human protein able to suppress growth of cancer cells and its
PT encoding polynucleotide -
PT
XX Claim 1; Page 14 (Disclosure); 38pp; Chinese.
PS
XX The present sequence represents human PP1030 protein, which has cancer
CC suppressing activity. The present invention describes a method for
CC the preparation of the protein by recombination, and the application
CC of the protein in treating diseases such as cancer.

XX SQ Sequence 283 AA;
Query Match 1.5%; Score 8; DB 23; Length 283;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 GPSLAAPA 31
Db 134 GPSLAAPA 141
RESULT 10
ID AAB10901 standard; Protein; 487 AA.
XX
XX AAB10901;
AC
XX 30-JAN-2001 (first entry)
DT
XX S. xyloosus DltA protein.
DE
XX dltA; dltB; dltC; dltD; dltABCD operon; antibacterial; D-alanine;
KW teichoic acid; endotoxin-like; inflammation; Gram-positive bacteria;
KW antimicrobial; D-alanine-D-alanyl carrier protein ligase; ds.
XX
XX Staphylococcus xyloosus.
OS
XX DE19912706-A1.
PN
XX 07-SEP-2000.
PD
XX 20-MAR-1999; 99DE-1012706.
PF
XX 05-MAR-1999; 99DE-1009636.
PR
XX (PETR-) PETRY GENMEDICS GMBH.
PA
XX Goetz F, Peschel A;
PI
XX WPI; 2000-588432/56.
DR
XX New staphylococcal DNA for dltABCD operons, useful e.g. for identifying
PT antibacterials and agents that reduce bacterial resistance to
PT antimicrobials -
XX
XX Claim 4; Fig 5; 20pp; German.
PS
XX This invention describes novel DNA sequences (A) of the dltABCD operon
CC from Staphylococcus xyloosus and Staphylococcus aureus. The products of
CC the invention have antibacterial activity. The proteins expressed by (A)
CC are involved in incorporation of D-alanine (Dala) into teichoic acid (I)
CC (which has endotoxin-like inflammatory activity) by Gram-positive
CC bacteria. Incorporation of Dala into (I) is correlated with sensitivity
CC of bacteria to antimicrobial agents, i.e. Dala is necessary for
CC resistance. When tested against wild-type S. aureus Sall3, the human
CC neutrophilic peptide defensin had minimum inhibitory concentration (MIC)
CC over 100 micro g/ml, but against a mutant in which the dltA gene has been
CC deleted it had MIC over 10 micro g/ml. Similar reductions in MIC were
CC determined for other cationic antibacterial peptides. (A), optionally
CC mutated, are used to study the function of their encoded proteins,
CC involved in incorporation of D-alanine (Dala) into teichoic acid (I).
CC Agents that reduce the inflammatory activity of (I) or incorporation of
CC Dala into (I) are used: (i) to increase the sensitivity of Gram-positive
CC bacteria to antimicrobial agents; (ii) to inhibit formation of biofilms
CC (particularly of staphylococci) on glass, metal or plastics surfaces
CC (e.g. catheters or cardiac pacemakers); and (iii) as antibacterials. This
CC sequence represents the Staphylococcus xyloosus D-alanine-D-alanyl carrier
CC protein ligase (dltA) which is described in the method of the invention.
XX
XX SQ Sequence 487 AA;
Query Match 1.5%; Score 8; DB 21; Length 487;

Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 409 LSLAENEE 416
Db 322 LSLAENEE 329
|||||

RESULT 11

AAW00292
ID AAW00292 standard; Peptide; 29 AA.

XX AC AAW00292;

XX DT 20-NOV-1996 (first entry)

XX DE Penicillin V amidohydrolase N-terminus.

XX KW Penicillin V amidohydrolase; PVA; F. oxysporum; strain 435;
XX KW hydrolysis; phenoxymethylpenicillin; 6-aminopenicillanic acid;
XX KW 6-APA; beta-lactam; semi-synthetic penicillin; expression vector;
XX KW recombinant production.

XX OS Fusarium oxysporum.

XX FH Key Location/Qualifiers

XX FT Misc-difference 7 /label= Ala, Lys

XX FT Misc-difference 9

XX FT /note= "Any amino acid"

XX FT Misc-difference 22

XX FT /label= Thr, Val

XX PN US5516679-A.

XX PD 14-MAY-1996.

XX PP 23-DEC-1994; 9AUS-0363475.

XX PR 23-DEC-1994; 9AUS-0363475.

XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX PI Burnett WV, Chiang S, Tonzi SM;

XX DR WPI; 1996-251011/25.

XX DNA encoding penicillin V amidohydrolase (PVA) from Fusarium
XX oxysporum - also recombinant vectors and host cells for production
XX of PVA for use in the manufacture of penicillin.

XX PS Claim 18; Fig 1; 46pp; English.

XX The sequences given in AAW00292-98 and AAW00300 are peptide fragments
XX derived from the secreted form of penicillin V amidohydrolase (PVA)
XX from F. oxysporum strain 435. The secreted form of PVA is a
XX glycoprotein of mol. wt. 65 kD. The seven amino acid fragment of
XX peptide C given in AAW00298 was used in the design of four probes by
XX reverse translation (see also AAT40247). This probe was used in the
XX identification of a PVA cDNA clone. PVA is used for the enzymatic
XX hydrolysis of penicillin V (phenoxymethylpenicillin) to 6-amino-
XX penicillanic acid (6-APA). 6-APA is the active beta-lactam nucleus
XX used in the manufacture of semi-synthetic penicillins. The PVA coding
XX sequences can be inserted into expression vectors for the recombinant
XX production of PVA in a suitable host, pref. Fusarium sp.

XX Sequence 29 AA;

Query Match 1.3%; Score 7; DB 17; Length 29;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 115 SLKLPGT 121

Db 15 SLKLPGT 21
|||||

RESULT 12

AAW97535
ID AAW97535 standard; peptide; 36 AA.

XX AC AAW97535;

XX DT 19-MAY-1999 (first entry)

XX DE Antigenic site of HN protein loop beta-5L01.

XX KW Antigenic site; haemagglutinin-neuraminidase; HN; paramyxoviridae;
XX KW virus epitope; attachment protein; vaccine; immunodominant epitope.

XX OS Simian virus 9.

XX PN WO9902695-A2.

XX PD 21-JAN-1999.

XX PF 08-JUL-1998; 98WO-NL00390.

XX PR 08-JUL-1997; 97EP-0202100.

XX PA (DIER-) STICHTING INST DIERHOUDERIJ EN DIERGEZON.

XX PI Langedijk JPM, Van Oirschot JT;

XX DR WPI; 1999-120896/10.

XX PT Isolated proteinaceous substance - comprising at least one virus
XX epitope derived from an attachment protein of a paramyxovirus

XX PS Disclosure; Page 49; 63pp; English.

XX AAW97452-571 represent antigenic sites derived from the
XX haemagglutinin-neuraminidase (HN) protein of the paramyxoviridae.
XX The specification describes 3-D models identifying a proteinaceous
XX substance comprising at least one virus epitope derived from the
XX attachment protein, which corresponds to an antigenic site present
XX one of the loops of HN. The antigenic sites can be used to produce
XX vaccines, to detect the viruses, and to select the immunodominant
XX epitope.

XX SQ Sequence 36 AA;

Query Match 1.3%; Score 7; DB 20; Length 36;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 108 QVPRPGT 114
Db 9 QVPRPGT 15
|||||

RESULT 13

ABB49653
ID ABB49653 standard; Protein; 59 AA.

XX AC ABB49653;

XX DT 05-FEB-2002 (first entry)

XX DE Listeria monocytogenes protein #2357.

XX KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX KW vitamin B12; bacterial infection; disease.

XX OS Listeria monocytogenes.

XX'

PN WO200177335-A2.
 XX 18-OCT-2001.
 PD 11-APR-2001; 2001WO-FR01118.
 PF 11-APR-2000; 2000FR-0004629.
 XX (INSP) INST PASTEUR.
 PA Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
 PI Dussurget O, Ghetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Krefit J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Maduenio E, De Pablos B, Wehlend J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 XX WPI; 2002-010914/01.
 DR Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment
 XX and prevention of *Listeria* and related bacterial infections, and
 PT related polypeptides
 PT
 XX Claim 6; SEQ ID No 2358; 192pp; French.
 PS The present invention relates to the genome sequence of *Listeria*
 XX *monocytogenes* EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in *L.*
 CC *monocytogenes* and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of *L. monocytogenes* and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate *L. monocytogenes*-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by *L.*
 CC *monocytogenes* and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 59 AA;
 Query Match 1.3%; Score 7; DB 23; Length 59;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 412 AENESEA 418
 Db 17 AENESEA 23
 RESULT 14
 AAB33405
 ID AAB33405 standard; Protein; 66 AA.
 AC AAB33405;
 XX 25-JAN-2001 (first entry)
 DT Pinus radiata transcription factor protein sequence #401.
 DE
 XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
 KW poplar; sweetgum; teak; mahogany; bz1p; G-box binding factor;
 KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
 KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
 KW type 2 Cys2His2; CCAAT box element; MYB.
 XX

OS Pinus radiata.
 XX WO200053724-A2.
 XX 14-SEP-2000.
 PD 09-MAR-2000; 2000WO-US06112.
 PF 11-MAR-1999; 99US-0266513.
 XX 18-AUG-1999; 99US-0149485.
 PR (GENE-) GENESIS RES & DEV CORP LTD.
 XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 PI Wood M, McGrath A, Shenk MA, Glenn M;
 XX WPI; 2000-579369/54.
 DR New isolated polynucleotide encoding a plant transcription factor for
 XX producing a plant e.g. a woody plant, preferably eucalyptus or pine,
 PT having modified gene expression or modified activity of a polypeptide
 PT
 XX Claim 8; Page 745; 747pp; English.
 PS The present invention relates to novel plant transcription factors from
 CC *Eucalyptus grandis* or *Pinus radiata*. The present sequence is one such a
 CC transcription factor. The transcription factor may be used to produce a
 CC plant having modified gene expression such as a woody plant e.g. a
 CC *eucalyptus*, pine, acacia, poplar, sweetgum, teak, or mahogany species or
 CC to modify the activity of a polypeptide in a plant. The transcription
 CC factors of the present invention are members from the following families
 CC of regulatory proteins: bz1p, bz1p family of G-box binding factors, basic
 CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
 CC zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2
 CC Cys2His2, CCAAT box elements and MYB.
 XX
 SQ Sequence 66 AA;
 Query Match 1.3%; Score 7; DB 21; Length 66;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 373 RLQWGI 379
 Db 53 RLQWGI 59
 RESULT 15
 ABP04937
 ID ABP04937 standard; Protein; 73 AA.
 XX AC ABP04937;
 XX 24-JUN-2002 (first entry)
 DT Human ORFX protein sequence SEQ ID NO:9856.
 DE
 XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 XX
 OS Homo sapiens.
 XX WO200192523-A2.
 PN 06-DEC-2001.
 PD


```
XX 21-MAR-2001; 2001WO-US09180.
PF
XX
XX 21-MAR-2000; 2000US-191078P.
PR
XX 23-MAY-2000; 2000US-206848P.
PR
XX 26-MAY-2000; 2000US-207727P.
PR
XX 23-OCT-2000; 2000US-242578P.
PR
XX 27-NOV-2000; 2000US-253625P.
PR
XX 22-DEC-2000; 2000US-257931P.
PR
XX 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR N-PSDB; AAS53316.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 11050; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, themselves and the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 75 AA;

Query Match 1.3%; Score 7; DB 22; Length 75;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 YFRRRKPF 52
DB 4 YFRRRKPF 10

RESULT 18
AAU38266
ID AAU38266 standard; Protein; 75 AA.
XX
XX AAU38266;
AC
XX
XX 14-FEB-2002 (first entry)
DT
XX
XX Salmonella typhi cellular proliferation protein #157.
DE
XX
XX Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
XX Salmonella typhi.
OS
XX WO200170955-A2.
XX
XX 27-SEP-2001.
PD

XX 21-MAR-2001; 2001WO-US09180.
PF
XX
XX 21-MAR-2000; 2000US-191078P.
PR
XX 23-MAY-2000; 2000US-206848P.
PR
XX 26-MAY-2000; 2000US-207727P.
PR
XX 23-OCT-2000; 2000US-242578P.
PR
XX 27-NOV-2000; 2000US-253625P.
PR
XX 22-DEC-2000; 2000US-257931P.
PR
XX 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR N-PSDB; AAS56125.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 13859; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, themselves and the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 75 AA;

Query Match 1.3%; Score 7; DB 22; Length 75;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 YFRRRKPF 52
DB 4 YFRRRKPF 10

RESULT 19
ABB16053
ID ABB16053 standard; Protein; 75 AA.
XX
XX ABB16053;
AC
XX
XX 23-JAN-2002 (first entry)
DT
XX
XX Human nervous system related polypeptide SEQ ID NO 4710.
DE
XX
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antiscikling; antianemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
```

OS	Homo sapiens.	PR	27-SEP-2000; 2000US-0235834.
XX		PR	27-SEP-2000; 2000US-0235836.
PN	WO200159063-A2.	PR	29-SEP-2000; 2000US-0236327.
XX		PR	29-SEP-2000; 2000US-0236367.
PD	16-AUG-2001.	PR	29-SEP-2000; 2000US-0236368.
XX		PR	29-SEP-2000; 2000US-0236369.
XX		PR	29-SEP-2000; 2000US-0236370.
XX		PR	02-OCT-2000; 2000US-0236802.
XX		PR	02-OCT-2000; 2000US-0237037.
PR	31-JAN-2000; 2000US-0179065.	PR	02-OCT-2000; 2000US-0237038.
PR	04-FEB-2000; 2000US-0180628.	PR	02-OCT-2000; 2000US-0237039.
PR	24-FEB-2000; 2000US-0184664.	PR	02-OCT-2000; 2000US-0237040.
PR	02-MAR-2000; 2000US-0186350.	PR	13-OCT-2000; 2000US-0239935.
PR	16-MAR-2000; 2000US-0189874.	PR	13-OCT-2000; 2000US-0239937.
PR	17-MAR-2000; 2000US-0190076.	PR	20-OCT-2000; 2000US-0240960.
PR	18-APR-2000; 2000US-0198123.	PR	20-OCT-2000; 2000US-0241785.
PR	19-MAY-2000; 2000US-0205515.	PR	20-OCT-2000; 2000US-0241786.
PR	07-JUN-2000; 2000US-0209467.	PR	20-OCT-2000; 2000US-0241787.
PR	28-JUN-2000; 2000US-0214886.	PR	20-OCT-2000; 2000US-0241808.
PR	30-JUN-2000; 2000US-0215135.	PR	20-OCT-2000; 2000US-0241809.
PR	07-JUL-2000; 2000US-0215647.	PR	20-OCT-2000; 2000US-0241826.
PR	07-JUL-2000; 2000US-0216880.	PR	20-OCT-2000; 2000US-0242221.
PR	11-JUL-2000; 2000US-0217487.	PR	01-NOV-2000; 2000US-0244617.
PR	11-JUL-2000; 2000US-0217496.	PR	08-NOV-2000; 2000US-0246474.
PR	14-JUL-2000; 2000US-0218290.	PR	08-NOV-2000; 2000US-0246475.
PR	26-JUL-2000; 2000US-0220963.	PR	08-NOV-2000; 2000US-0246476.
PR	26-JUL-2000; 2000US-0220964.	PR	08-NOV-2000; 2000US-0246477.
PR	14-AUG-2000; 2000US-0220964.	PR	08-NOV-2000; 2000US-0246478.
PR	14-AUG-2000; 2000US-0224519.	PR	08-NOV-2000; 2000US-0246523.
PR	14-AUG-2000; 2000US-0225213.	PR	08-NOV-2000; 2000US-0246524.
PR	14-AUG-2000; 2000US-0225214.	PR	08-NOV-2000; 2000US-0246525.
PR	14-AUG-2000; 2000US-0225266.	PR	08-NOV-2000; 2000US-0246526.
PR	14-AUG-2000; 2000US-0225267.	PR	08-NOV-2000; 2000US-0246527.
PR	14-AUG-2000; 2000US-0225268.	PR	08-NOV-2000; 2000US-0246528.
PR	14-AUG-2000; 2000US-0225270.	PR	08-NOV-2000; 2000US-0246532.
PR	14-AUG-2000; 2000US-0225447.	PR	08-NOV-2000; 2000US-0246609.
PR	14-AUG-2000; 2000US-0225757.	PR	08-NOV-2000; 2000US-0246610.
PR	14-AUG-2000; 2000US-0225758.	PR	08-NOV-2000; 2000US-0246611.
PR	18-AUG-2000; 2000US-0225759.	PR	08-NOV-2000; 2000US-0246613.
PR	18-AUG-2000; 2000US-0226279.	PR	17-NOV-2000; 2000US-0249207.
PR	22-AUG-2000; 2000US-0226681.	PR	17-NOV-2000; 2000US-0249208.
PR	22-AUG-2000; 2000US-0226686.	PR	17-NOV-2000; 2000US-0249209.
PR	22-AUG-2000; 2000US-0227182.	PR	17-NOV-2000; 2000US-0249210.
PR	23-AUG-2000; 2000US-0227009.	PR	17-NOV-2000; 2000US-0249211.
PR	30-AUG-2000; 2000US-0228924.	PR	17-NOV-2000; 2000US-0249212.
PR	01-SEP-2000; 2000US-0229287.	PR	17-NOV-2000; 2000US-0249213.
PR	01-SEP-2000; 2000US-0229343.	PR	17-NOV-2000; 2000US-0249214.
PR	01-SEP-2000; 2000US-0229344.	PR	17-NOV-2000; 2000US-0249215.
PR	01-SEP-2000; 2000US-0229345.	PR	17-NOV-2000; 2000US-0249216.
PR	05-SEP-2000; 2000US-0229509.	PR	17-NOV-2000; 2000US-0249217.
PR	05-SEP-2000; 2000US-0229513.	PR	17-NOV-2000; 2000US-0249218.
PR	06-SEP-2000; 2000US-0230437.	PR	17-NOV-2000; 2000US-0249244.
PR	06-SEP-2000; 2000US-0230438.	PR	17-NOV-2000; 2000US-0249245.
PR	08-SEP-2000; 2000US-0231242.	PR	17-NOV-2000; 2000US-0249264.
PR	08-SEP-2000; 2000US-0231243.	PR	17-NOV-2000; 2000US-0249265.
PR	08-SEP-2000; 2000US-0231244.	PR	17-NOV-2000; 2000US-0249297.
PR	08-SEP-2000; 2000US-0231413.	PR	17-NOV-2000; 2000US-0249299.
PR	08-SEP-2000; 2000US-0231414.	PR	17-NOV-2000; 2000US-0249300.
PR	08-SEP-2000; 2000US-0232080.	PR	01-DEC-2000; 2000US-0250391.
PR	08-SEP-2000; 2000US-0232081.	PR	01-DEC-2000; 2000US-0251160.
PR	12-SEP-2000; 2000US-0231968.	PR	05-DEC-2000; 2000US-0251030.
PR	14-SEP-2000; 2000US-0232197.	PR	05-DEC-2000; 2000US-0251988.
PR	14-SEP-2000; 2000US-0232198.	PR	05-DEC-2000; 2000US-0256719.
PR	14-SEP-2000; 2000US-0232399.	PR	06-DEC-2000; 2000US-0251479.
PR	14-SEP-2000; 2000US-0232400.	PR	08-DEC-2000; 2000US-0251856.
PR	14-SEP-2000; 2000US-0232401.	PR	08-DEC-2000; 2000US-0251868.
PR	14-SEP-2000; 2000US-0233063.	PR	08-DEC-2000; 2000US-0251869.
PR	14-SEP-2000; 2000US-0233064.	PR	08-DEC-2000; 2000US-0251989.
PR	14-SEP-2000; 2000US-0233065.	PR	08-DEC-2000; 2000US-0251990.
PR	21-SEP-2000; 2000US-0234223.	PR	11-DEC-2000; 2000US-0254097.
PR	21-SEP-2000; 2000US-0234274.	PR	05-JAN-2001; 2001US-0259678.
PR	25-SEP-2000; 2000US-0234997.	XX	
PR	25-SEP-2000; 2000US-0234998.	XX	
PR	26-SEP-2000; 2000US-0235484.	PA	(HUMA-) HUMAN GENOME SCI INC.

```

XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-541565/60.
XX N-PSDB; ABA12379.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
XX Claim 11; SEQ ID NO 4710; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-ABA21534) and proteins
CC (AB114678-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 75 AA;
SQ
Query Match 1.3%; Score 7; DB 22; Length 75;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 PRPGTSL 116
DB 46 PRPGTSL 52
|||||

RESULT 20
AAU42302
ID AAU42302 standard; Protein; 82 AA.
XX
XX AAU42302;
XX
XX 27-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #3198.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
OS
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US12865.
XX
XX 21-APR-2000; 2000US-199047P.
XX
XX 02-JUN-2000; 2000US-208841P.
XX
XX 07-JUL-2000; 2000US-216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

```

```

XX WPI; 2001-616774/71.
XX N-PSDB; AAS59516.
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Example 1; SEQ ID NO 3497; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 82 AA;
SQ
Query Match 1.3%; Score 7; DB 22; Length 82;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 LRPSTQS 151
DB 56 LRPSTQS 62
|||||

RESULT 21
ABP35330
ID ABP35330 standard; Protein; 85 AA.
XX
XX ABP35330;
XX
XX 08-JUL-2002 (first entry)
XX
XX Human ORF4303 protein, SEQ ID NO:8606.
XX
XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibitor; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neurotropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX
XX Homo sapiens.
OS
XX
XX WO200190366-A2.
XX
XX 29-NOV-2001.
XX
XX 24-MAY-2001; 2001WO-US17076.
XX

```

```

XX 24-MAY-2000; 2000US-206690P.
XX (CURA-) CURAGEN CORP.
XX Leach MD, Shimkets RA;
XX WPI; 2002-106200/14.
XX N-PSDB; ABN79356.
XX Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and disorders related to organ
XX transplantation
XX Claim 10; Page 2394-2395; 2508pp; English.
XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
XX designated ORF (open reading frame) 1-4534, and sequences ABN75054-
XX ABN79587 represent cDNAs encoding them. The invention also encompasses
XX polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
XX referred to as ORFX) proteins, polynucleotides at least 85% identical to
XX the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
XX polynucleotides, the recombinant production of ORFX proteins, antibodies
XX specific for ORFX proteins, methods of detecting ORFX polynucleotides and
XX polypeptides, methods of screening for modulators of ORFX expression or
XX activity, and methods of screening individuals for a predisposition to an
XX ORFX-associated disorder. The ORFX proteins of the invention have a wide
XX range of biological activities, such as cytokine, cell proliferation,
XX cell differentiation, immune modulation, haematopoiesis regulation,
XX tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
XX chemokinetic activity, haemostatic activity, thrombolytic activity,
XX receptor/ligand, antiinflammatory activity, tumour inhibition activity,
XX and antiinfective activity, and may also be involved in the determination
XX of bodily characteristics, fertility and behaviour. ORFX proteins,
XX nucleic acids and antibodies may be used in the treatment of cancers,
XX other proliferative disorders such as psoriasis and benign tumours,
XX neurological disorders such as epilepsy and Alzheimer's disease,
XX cardiovascular diseases, immune system disorders, disorders related to
XX organ transplantation, disorders of tissue growth and regeneration,
XX diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
XX storage disease, and infectious diseases caused by viral, bacterial,
XX fungal and other pathogens. ORFX nucleic acids may also be used as a
XX source of primers and probes, in the detection of ORFX genomic sequences
XX or transcripts, in the identification and cloning of homologous
XX sequences, in genetic diagnosis, and in forensic biology. The ORFX
XX nucleic acids may additionally be used to produce transgenic animals
XX which may be useful for studying the function and/or activity of ORFX
XX protein, and in drug screening. The ORFX proteins may also be used as
XX immunogens to generate specific antibodies, which are useful in the
XX diagnosis, treatment and monitoring of ORFX-associated diseases.
XX SQ Sequence 85 AA;
    Query Match 1.3%; Score 7; DB 23; Length 85;
    Best Local Similarity 100.0%; Pred. No. 1e+02;
    Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 150 QSGRPGT 156
Db 20 QSGRPGT 26
    |||||
    |||||

RESULT 22
AAV19803
ID AAV19803 standard; Protein; 103 AA.
XX AC AAV19803;
XX 19-JUL-1999 (first entry)
DT 19-JUL-1999 (first entry)
DE B. burgdorferi antigenic protein, t924.aa.
XX Antigenic protein; vaccine; Lyme disease; infection; detection.
XX Borrelia burgdorferi.
XX WO9859071-A1.
XX 30-DEC-1998.
XX 18-JUN-1998; 98WO-US12718.
XX 03-SEP-1997; 97US-0057483.
XX 20-JUN-1997; 97US-0050359.
XX 22-JUL-1997; 97US-0053344.
XX 22-JUL-1997; 97US-0053377.
XX B. burgdorferi antigenic protein, t924.aa.
XX

```

```

KW Antigenic protein; vaccine; Lyme disease; infection; detection.
XX Borrelia burgdorferi.
XX WO9859071-A1.
XX 30-DEC-1998.
XX 18-JUN-1998; 98WO-US12718.
XX 03-SEP-1997; 97US-0057483.
XX 20-JUN-1997; 97US-0050359.
XX 22-JUL-1997; 97US-0053344.
XX 22-JUL-1997; 97US-0053377.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (MEDI-) MEDIMUNE INC.
XX Choi GH, Erwin AL, Hanson MS, Lathigra R;
XX WPI; 1999-189980/16.
XX N-PSDB; AAX61500.
XX New isolated Borrelia burgdorferi nucleic acids - used to develop
XX products for the diagnosis, prevention and treatment of diseases
XX caused by Borrelia, particularly Lyme disease
XX Claim 12; Page 70; 275pp; English.
XX This sequence represents a Borrelia burgdorferi (Bb) protein of the
XX invention, which is suitable for use in a vaccine. The Bb polypeptides
XX can be used in vaccines for eliciting protective antibodies to members of
XX the Borrelia genus, particularly for the use against Lyme disease in
XX humans and animals. They can be used for preventing or attenuating an
XX infection caused by a member of the Borrelia genus. The products can also
XX be used for detection of members of the Borrelia genus.
XX SQ Sequence 103 AA;
    Query Match 1.3%; Score 7; DB 20; Length 103;
    Best Local Similarity 100.0%; Pred. No. 1.2e+02;
    Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 TOMLEKS 66
Db 1 TOMLEKS 7
    |||||
    |||||

RESULT 23
AAV19802
ID AAV19802 standard; Protein; 133 AA.
XX AC AAV19802;
XX 19-JUL-1999 (first entry)
DT 19-JUL-1999 (first entry)
DE B. burgdorferi antigenic protein, f924.aa.
XX Antigenic protein; vaccine; Lyme disease; infection; detection.
XX Borrelia burgdorferi.
XX WO9859071-A1.
XX 30-DEC-1998.
XX 18-JUN-1998; 98WO-US12718.
XX 03-SEP-1997; 97US-0057483.
XX 20-JUN-1997; 97US-0050359.
XX 22-JUL-1997; 97US-0053344.
XX 22-JUL-1997; 97US-0053377.
XX

```

PA (HUMA-) HUMAN GENOME SCI INC.
 PA (MEDI-) MEDIMMUNE INC.
 XX Choi GH, Erwin AL, Hanson MS, Lathigra R;
 XX WPI: 1999-189980/16.
 DR N-PSDB; AAX61499.
 XX
 PT New isolated *Borrelia burgdorferi* nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of diseases
 PT caused by *Borrelia*, particularly Lyme disease
 XX
 PS Claim 12; Page 70; 275pp; English.
 XX
 CC This sequence represents a *Borrelia burgdorferi* (Bb) protein of the
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides
 CC can be used in vaccines for eliciting protective antibodies to members of
 CC the *Borrelia* genus, particularly for the use against Lyme disease in
 CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the *Borrelia* genus. The products can also
 CC be used for detection of members of the *Borrelia* genus.
 XX
 SQ Sequence 133 AA;
 Query Match 1.3%; Score 7; DB 20; Length 133;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 60 TOMLEKS 66
 |||||
 Db 31 TOMLEKS 37
 RESULT 24
 AAG26620
 ID AAG26620 standard; Protein; 144 AA.
 XX
 AC AAG26620;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 31144.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0121180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.

```
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 7; DB 21; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 413 ENEEEA 419
Db 112 ENEEEA 118

RESULT 25
AAG48630
ID AAG48630 standard; Protein; 144 AA.
XX
AC AAG48630;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 61432.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
```

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 7; DB 21; Length 144;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 413 ENEEEAA 419
 |||||
 Db 112 ENEEEAA 118

RESULT 26
 AAB33260
 ID AAB33260 standard; Protein; 146 AA.
 XX
 AC AAB33260;
 XX
 DT 25-JAN-2001 (first entry)
 XX
 DE Pinus radiata transcription factor protein sequence #306.
 XX
 KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
 KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
 KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
 KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
 KW type 2 Cys2His2; CCAAT box element; MYB.
 XX
 OS Pinus radiata.
 XX
 PN WO200053724-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06112.
 XX
 PR 11-MAR-1999; 99US-0266513.
 PR 18-AUG-1999; 99US-0149485.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 PI Wood M, McGrath A, Shenk MA, Glenn M;
 XX
 WPI; 2000-579369/54.
 DR
 XX
 PT New isolated polynucleotide encoding a plant transcription factor for
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
 PT having modified gene expression or modified activity of a polypeptide
 PT -
 XX
 PS Claim 8; Page 698; 747pp; English.
 XX
 CC The present invention relates to novel plant transcription factors from
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
 CC transcription factor. The transcription factor may be used to produce a
 CC plant having modified gene expression such as a woody plant e.g. a
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
 CC to modify the activity of a polypeptide in a plant. The transcription
 CC factors of the present invention are members from the following families
 CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
 CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
 CC zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2
 CC Cys2His2, CCAAT box elements and MYB.
 XX
 SQ Sequence 146 AA;

Query Match 1.3%; Score 7; DB 21; Length 146;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 373 RLQMG1 379
 |||||
 Db 97 RLQMG1 103

RESULT 27
 ABB60190
 ID ABB60190 standard; Protein; 171 AA.
 XX
 AC ABB60190;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 7362.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 WPI; 2001-656860/75.
 DR N-PSDB; ABL04293.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 7362; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 171 AA;

Query Match 1.3%; Score 7; DB 22; Length 171;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 413 ENEEEAA 419
 |||||
 Db 71 ENEEEAA 77

RESULT 28
 ABB71753
 ID ABB71753 standard; Protein; 209 AA.

[illegible]

```
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146708.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 7; DB 21; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 413 ENEEAA 419
Db 181 ENEEAA 187
|||||

RESULT 30
AAG48629
ID AAG48629 standard; Protein; 213 AA.
XX AC AAG48629;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 61431.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
```

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128645.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155119.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 26-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161921.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 7; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 413 ENEEEAA 419
 |||||
 Db 181 ENEEEAA 187

RESULT 31
 AAR60575
 ID AAR60575 standard; Protein; 215 AA.

AC AAR60575;
 DT 25-MAR-2003 (updated)
 DT 01-APR-1995 (first entry)
 XX
 DE House dust mite allergen DerpVII cDNA.
 XX
 KW DerpVII allergen; antiallergic; allergy diagnosis.
 XX
 OS Dermatophagoides pteronyssinus.
 XX

PN W09420614-A1.

PD 15-SEP-1994.

XX
 PF 11-MAR-1994; 94WO-AU00117.
 XX
 PR 12-MAR-1993; 93US-0031141.
 PR 22-JUN-1993; 93US-0081540.

XX (CHIL-) INST CHILD HEALTH RES.

XX Chua K, Thomas WR;

XX WPI; 1994-303021/37.

DR N-PSDB; AAQ71400.

XX New nucleic acid encoding specific dust mite allergens - and
 PT related vectors, transformed cells, peptides and antibodies,
 PT useful for desensitisation and diagnosis.
 XX Claim 7; Page 36-37; 67pp; English.
 PS DerpVII antigen is useful as antiallergic reagent for treating
 CC sensitivity to house dust mite allergens. The DNA can be used
 CC as a probe to detect the sensitivity of an individual to the
 CC allergen.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX

SQ Sequence 215 AA;

Query Match 1.3%; Score 7; DB 15; Length 215;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 400 MTLTSPE 406
 |||||
 Db 152 MTLTSPE 158

RESULT 32

AAAY25586

ID AAY25586 standard; protein; 215 AA.

XX AC AAY25586;

XX DT 30-SEP-1999 (first entry)

XX DE D. pteronyssinus allergen Der p 7 protein fragment.

XX Major histocompatibility complex; class II; desensitising; human;
 KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; stinging;
 KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KW mite; gerbil; vaccine; treatment; prevention; hypersensitivity.

XX Dermatophagoides pteronyssinus.

XX W09934826-A1.

XX PD 15-JUL-1999.

XX PF 11-JAN-1999; 99WO-GB00080.

XX PR 21-SEP-1998; 98GB-0020474.

XX PR 09-JAN-1998; 98GB-0000445.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Kay AB, Larche M;

XX WPI; 1999-458255/38.

XX Desensitizing patients to polypeptide allergens

XX Example 6; Page 51; 117pp; English.

XX This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitising patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of

CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents the house dust mite (Dermatophagoides pteronyssinus) allergen
 CC Der p 7.
 XX
 XX

SQ Sequence 215 AA;

Query Match 1.3%; Score 7; DB 20; Length 215;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 400 MTLTSFE 406
 DB 152 MTLTSFE 158
 |||||

RESULT 33

AAW81726
 ID AAW81726 standard; Protein; 267 AA.

XX
 AC AAW81726;

XX
 DT 27-JAN-1999 (first entry)

XX
 DE M. tuberculosis immunogenic polypeptide TBH-29.

XX
 KW Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
 KW vaccine; pharmaceutical; infection; diagnosis.

XX
 OS Mycobacterium tuberculosis.

XX
 PN WO9816646-A2.

XX
 PD 23-APR-1998.

XX
 PF 07-OCT-1997; 97WO-US18293.

XX
 PR 13-MAR-1997; 97US-0818112.

XX
 PR 11-OCT-1996; 96US-0730510.

XX
 PA (CORI-) CORIXA CORP.

XX
 PI Campos-Neto A. Dillon DC, Houghton R. Lodes MJ;
 PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;

XX
 DR WPI; 1998-261042/23.

XX
 DR N-PSDB; AAV64512.

XX
 PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used
 PT to develop products for the detection of M. tuberculosis infection
 PT and for diagnosis, treatment and prevention of tuberculosis

XX
 PS Example 3b; Page 149-150; 230pp; English.

XX
 CC This sequence represents an immunogenic portion of a soluble
 CC Mycobacterium tuberculosis (MT) antigen which can be used in a method
 CC for inducing protective immunity against tuberculosis (TB). This
 CC sequence can be formulated into vaccines and/or pharmaceutical
 CC compositions for immunising against M. tuberculosis infection or may
 CC be used for the diagnosis of tuberculosis.

XX
 SQ Sequence 267 AA;

Query Match 1.3%; Score 7; DB 19; Length 267;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 VPRPGTS 115
 DB 78 VPRPGTS 84
 |||||

RESULT 34

AAW64359

ID AAW64359 standard; Protein; 267 AA.

XX
 AC AAW64359;

XX
 DT 09-NOV-1998 (first entry)

XX
 DE Mycobacterium tuberculosis antigen TBH-29.

XX
 KW Tuberculosis; infection; diagnosis; antigen; TBH-29.

XX
 OS Mycobacterium tuberculosis strain H37Rv.

XX
 PN WO9816645-A2.

XX
 PD 23-APR-1998.

XX
 PF 07-OCT-1997; 97WO-US18214.

XX
 PR 13-MAR-1997; 97US-0818111.

XX
 PR 11-OCT-1996; 96US-0729622.

XX
 PA (CORI-) CORIXA CORP.

XX
 PI Campos-Neto A. Dillon DC, Houghton R. Lodes MJ;
 PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;

XX
 DR WPI; 1998-251292/22.

XX
 DR N-PSDB; AAV44403.

XX
 PT New isolated Mycobacterium tuberculosis polypeptides and DNA - used
 PT to develop products for the detection of M. tuberculosis infection
 PT and diagnosis of tuberculosis

XX
 PS Example 3; Page 157-158; 250pp; English.

XX
 CC This polypeptide comprises an antigenic portion of Mycobacterium
 CC tuberculosis antigen TBH-29. A DNA molecule (see AAV44403) encoding
 CC the polypeptide was isolated from a M. tuberculosis strain H37Rv
 CC genomic library. The invention relates to compositions and methods
 CC for diagnosing tuberculosis. It provides polypeptides (see
 CC AAW64291-W64379) comprising an antigenic portion of a soluble M.
 CC tuberculosis antigen, or an immunogenic portion of an M.
 CC tuberculosis antigen, as well as DNA sequences encoding such
 CC polypeptides, recombinant expression vectors and transformed or
 CC transfected host cells. Also claimed are methods and diagnostic
 CC kits for detecting M. tuberculosis infection in a patient using
 CC these polypeptides, antibodies or oligonucleotide probes and
 CC primers, for the diagnosis of tuberculosis.

XX
 SQ Sequence 267 AA;

Query Match 1.3%; Score 7; DB 19; Length 267;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 VPRPGTS 115
 DB 78 VPRPGTS 84
 |||||

RESULT 35

AAV39156

ID AAV39156 standard; Protein; 267 AA.

XX
 AC AAV39156;

XX
 DT 05-NOV-1999 (first entry)

XX
 DE M. tuberculosis antigen TBH-29 amino acid sequence.

XX

KW Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
 KW immunotherapy; diagnosis; immunisation; vaccine; infection;
 KW immune response; skin test.

XX Mycobacterium tuberculosis.

XX WO9942076-A2.

XX 26-AUG-1999.

XX 17-FEB-1999; 99WO-US03268.

XX 05-MAY-1998; 98US-0072967.

XX 18-FEB-1998; 98US-0025197.

XX (CORI-) CORIXA CORP.

XX Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
 PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;

XX WPI; 1999-527409/44.

XX N-PSDB; AAZ19313.

XX New antigens from Mycobacterium tuberculosis useful in diagnostic
 PT skin tests and protective or therapeutic vaccines or compositions

XX Example 3; Page 144; 299pp; English.

XX The present invention describes polypeptides comprising an immunogenic
 CC part of a Mycobacterium tuberculosis antigen (Ag). Also described
 CC are vaccines and fusion protein containing M. tuberculosis Ag's.
 CC M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and
 CC other polypeptides fragments, can be used in pharmaceutical compositions
 CC or vaccines to generate a protective or therapeutic immune response to
 CC M. tuberculosis and as reagents in skin tests for diagnosis of
 CC tuberculosis. Ag can induce proliferation of, or cytokine secretion
 CC by, T, B or natural killer cells and/or macrophages in
 CC tuberculosis-immune subjects. AAZ19249 to AAZ19460 and AAY39083 to
 CC AAY39225 are used in the exemplification of the present invention.

XX Sequence 267 AA;

Query Match 1.3%; Score 7; DB 20; Length 267;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 VPRPGTS 115

Db 78 VPRPGTS 84

RESULT 36

AAV39013
 ID AAV39013 standard; Protein; 267 AA.

XX AAV39013;

XX 05-NOV-1999 (first entry)

XX M. tuberculosis recombinant antigen protein TbH-29.

XX Antigen; diagnosis; detection; infection; antibody; immunisation;
 KW vaccine; immunity.

XX Mycobacterium tuberculosis.

XX WO9942118-A2.

XX 26-AUG-1999.

XX 17-FEB-1999; 99WO-US03265.

XX 05-MAY-1998; 98US-0072596.

PR 18-FEB-1998; 98US-0024753.
 XX (CORI-) CORIXA CORP.
 XX Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
 PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;

XX WPI; 1999-527416/44.

XX N-PSDB; AAZ19101.

XX New polypeptide comprising antigenic portions of M. tuberculosis

XX Example 2; Page 189-190; 323pp; English.

XX This invention describes novel recombinant antigens and their encoding
 CC nucleic acids derived from Mycobacterium tuberculosis. The novel in a
 CC polypeptides are useful for detecting M. tuberculosis infection in a
 CC biological sample by detecting antibodies which bind with the
 CC polypeptides, and are useful as vaccines for immunizing against
 CC M. tuberculosis infection. The new detection methods are needed as
 CC current vaccination strategies do not provide 100% immunity.

XX Sequence 267 AA;

Query Match 1.3%; Score 7; DB 20; Length 267;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 VPRPGTS 115

Db 78 VPRPGTS 84

RESULT 37

ABG07558
 ID ABG07558 standard; Protein; 287 AA.

XX ABG07558;

XX 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #7549.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS71745.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

XX Claim 20; SEQ ID No 37917; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAG0010-ABG3077 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 287 AA;

Query Match 1.3%; Score 7; DB 22; Length 287;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 QPVTALN 301

Db 224 QPVTALN 230

RESULT 38

AAG23634
 ID AAG23634 standard; Protein; 290 AA.

AC AAG23634;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 27015.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-01451192.
PR 23-JUL-1999; 99US-01451145.
PR 23-JUL-1999; 99US-01452118.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-01452176.
PR 27-JUL-1999; 99US-01452913.
PR 27-JUL-1999; 99US-01459118.
PR 27-JUL-1999; 99US-01459119.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.38; Score 7; DB 21; Length 290;

Best Local Similarity 100.0%; Pred. No. 3.1e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 24 GPSLAAP 30
Db 36 GPSLAAP 42
|||||

RESULT 39

AAG23633
ID AAG23633 standard; Protein; 307 AA.

XX AC AAG23633;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 27014.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

 Query Match 1.3%; Score 7; DB 21; Length 307;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 24 GPSLAAP 30
 Db 53 GPSLAAP 59

 RESULT 40
 AAY74359
 ID AAY74359 standard; Protein; 309 AA.
 XX AC AAY74359;
 XX DT 21-MAR-2000 (first entry)
 XX DE Neisseria gonorrhoeae ORF 047 protein sequence SEQ ID NO:194.
 XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy.
 XX OS Neisseria gonorrhoeae.
 XX PN WO9957280-A2.
 XX PD 11-NOV-1999.
 XX PF 30-APR-1999; 99WO-US09346.
 XX PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX DR WPI; 2000-062150/05.
 DR N-PSDB; AAZ53121.
 XX PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 XX PS Claim 2; Page 246; 1453pp; English.
 XX CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC the invention can be used as vaccines, antibodies and compositions of
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC presence of Neisseria bacteria (e.g. meningitis and septicaemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also

CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX SQ Sequence 309 AA;

 Query Match 1.3%; Score 7; DB 21; Length 309;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 85 YIDEIDV 91
 Db 141 YIDEIDV 147

 RESULT 41
 AAY74361
 ID AAY74361 standard; Protein; 309 AA.
 XX AC AAY74361;
 XX DT 21-MAR-2000 (first entry)
 XX DE Neisseria meningitidis ORF 047 protein sequence SEQ ID NO:198.
 XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy.
 XX OS Neisseria meningitidis.
 XX PN WO9957280-A2.
 XX PD 11-NOV-1999.
 XX PF 30-APR-1999; 99WO-US09346.
 XX PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX DR WPI; 2000-062150/05.
 DR N-PSDB; AAZ53123.
 XX PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 XX PS Claim 2; Page 248; 1453pp; English.
 XX CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC the invention can be used as vaccines, antibodies and compositions of
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC presence of Neisseria bacteria (e.g. meningitis and septicaemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also

CC may also be used in gene therapy protocols.

XX Sequence 309 AA;

Query Match 1.3%; Score 7; DB 21; Length 309;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 85 YIDEIDV 91

|||||

Db 141 YIDEIDV 147

RESULT 42

AAG23632

ID AAG23632 standard; Protein; 310 AA.

XX

AC AAG23632;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 27013.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR

PR 05-MAR-1999; 99US-0123180.

PR

PR 09-MAR-1999; 99US-0123548.

PR

PR 23-MAR-1999; 99US-0125788.

PR

PR 25-MAR-1999; 99US-0126264.

PR

PR 29-MAR-1999; 99US-0126785.

PR

PR 01-APR-1999; 99US-0127462.

PR

PR 06-APR-1999; 99US-0128234.

PR

PR 08-APR-1999; 99US-0128714.

PR

PR 16-APR-1999; 99US-0129845.

PR

PR 19-APR-1999; 99US-0130077.

PR

PR 21-APR-1999; 99US-0130449.

PR

PR 23-APR-1999; 99US-0130510.

PR

PR 23-APR-1999; 99US-0130891.

PR

PR 28-APR-1999; 99US-0131449.

PR

PR 30-APR-1999; 99US-0132048.

PR

PR 30-APR-1999; 99US-0132407.

PR

PR 04-MAY-1999; 99US-0132484.

PR

PR 05-MAY-1999; 99US-0132485.

PR

PR 06-MAY-1999; 99US-0132486.

PR

PR 06-MAY-1999; 99US-0132487.

PR

PR 07-MAY-1999; 99US-0132863.

```
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155131.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159322.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 7; DB 21; Length 310;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

06-AUG-1999; 24 GPSLAAP 30
Db 56 GPSLAAP 62

RESULT 43
AAY74360
ID AAY74360 standard; Protein; 312 AA.
XX
AC AAY74360;
XX
XX 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 047 protein sequence SEQ ID NO:196.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX
FN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
Tettelin H, Venter JC;
XX
XX WPI; 2000-062150/05.
DR N-PSDB; AAZ53122.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
vaccines and diagnostics
XX
PS Claim 2; Page 246; 1453pp; English.
XX
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
PCR primers used in the exemplification of the present invention. The
polypeptides, the polynucleotides, antibodies and compositions of
the invention can be used as vaccines, as diagnostic reagents, and as
immunogenic compositions. The polypeptides can be used in the
manufacture of medicaments for treating or preventing infection due to
Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
presence of Neisseria bacteria, or to raise antibodies. They may also
be used to screen for agonists or antagonists, which may themselves
have use as antibacterial agents. The polynucleotides of the invention
may also be used in gene therapy protocols.
SQ Sequence 312 AA;

Query Match 1.3%; Score 7; DB 21; Length 312;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 144 YIDEIDV 150
|||||

RESULT 44

AAV58367
ID AAV58367 standard; Protein; 315 AA.

XX
AC AAV58367;

XX
DT 27-MAR-2000 (first entry)

XX
DE Anabaena variabilis AavI restriction endonuclease (AavI).

XX
KW Restriction endonuclease; AavI; restriction modification system;
KW AavI; Escherichia coli; recombinant expression.

XX
OS Anabaena variabilis.

XX
FN US6004793-A.

XX
PD 21-DEC-1999.

XX
PF 18-AUG-1998; 98US-0135639.

XX
PR 18-AUG-1998; 98US-0135639.

XX
PA (NEWE) NEW ENGLAND BIOLABS INC.

XX
PI Dalton MA, Wilson GG, Xu S, Lunnen KD;

XX
DR WPI; 2000-085793/07.

DR
N-PSDB; AAZ55711.

XX
PT Recombinant Escherichia coli engineered to express the restriction
PT endonuclease AavI -

XX
PS Example 1; Fig 3; 18pp; English.

XX
CC This sequence represents the Anabaena variabilis AavI restriction
CC endonuclease (AavI). This enzyme, along with the corresponding
CC AavI modification methylase (AAV58366), are the components of the
CC Anabaena variabilis restriction modification (R-M) system. This system
CC enable bacteria to destroy foreign DNA, while being able to protect
CC their own DNA and distinguish it from foreign DNA. Both the AavI
CC restriction endonuclease and the modification methylase recognise the
CC same DNA sequence, 5'-CYCGRG-3', the endonuclease cleaving between the
CC first and second bases on both strands, and the methylase methylating a
CC particular nucleotide in the sequence. Following methylation, the
CC recognition sequence is no longer cleaved by the cognate restriction
CC endonuclease. The DNA of the cell is fully modified by the methylase,
CC and is therefore completely insensitive to the presence of the
CC endogenous restriction endonuclease, whereas only unmodified,
CC identifiably foreign DNA is susceptible to endonuclease recognition and
CC cleavage. The invention relates to recombinant Escherichia coli (ATCC
CC 98837) engineered to express the restriction endonuclease AavI. Prior to
CC transformation with the AavI-M genes, the host cells were first
CC transformed with a plasmid encoding the isoschizomer BsoBI methylase,
CC enabling pre-modification of E. coli DNA against AavI endonuclease
CC digestion. The transformed E. coli may be used to recombinantly produce
CC the restriction endonuclease AavI in large quantities, improving enzyme
CC yield and simplifying enzyme purification. Restriction endonucleases are
CC useful tools for creating and analysing DNA molecules in the laboratory.
CC For example, they may be used in medical diagnostic applications such as
CC cancer gene mutation detection and pathogen detection via PCR.

XX
SQ Sequence 315 AA;

Query Match

Best Local Similarity 1.3%; Score 7; DB 21; Length 315;
ID ABG20636 standard; Protein; 315 AA.

XX
AC AAV58367;

XX
QY 469 VEQARAL 475

Db 39 VEQARAL 45
|||||

RESULT 45

ABB57333
ID ABB57333 standard; Protein; 317 AA.

XX
AC ABB57333;

XX
DT 07-MAR-2002 (first entry)

XX
DE Mouse ischaemic condition related protein sequence SEQ ID NO:928.

XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.

XX
OS Mus musculus.

XX
FN WO200188188-A2.

XX
PD 22-NOV-2001.

XX
PF 18-MAY-2001; 2001WO-JP04192.

XX
PR 18-MAY-2000; 2000JP-0145977.

XX
PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

XX
DR WPI; 2002-034733/04.

DR
N-PSDB; ABI99815.

XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -

XX
PS Claim 2; Page 2336-2338; 2690pp; English.

XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.

XX
SQ Sequence 317 AA;

Query Match

Best Local Similarity 1.3%; Score 7; DB 23; Length 317;
ID ABG20636 standard; Protein; 321 AA.

XX
AC ABB57333;

XX
QY 172 PITSSSG 178

XX
DB 277 PITSSSG 283

RESULT 46

ABG20636
ID ABG20636 standard; Protein; 321 AA.

XX
AC ABG20636;

QY


```

XX SQ Sequence 340 AA;
Query Match 1.3%; Score 7; DB 21; Length 340;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 272 SALKQOE 278
Db 4 SALKQOE 10
|||||

RESULT 51
AAB96616
ID AAB96616 standard; Protein; 363 AA.
XX
AC AAB96616;
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative P. abyssi protein #20.
XX
KW Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS Pyrococcus abyssi.
XX
PN FR2792651-A1.
XX
PD 27-OCT-2000.
XX
PF 21-APR-1999; 99FR-0005034.
XX
PR 21-APR-1999; 99FR-0005034.
XX
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
PI Querellou J, Weissenbach J, Saurin W, Heilig R;
XX
DR WPI; 2001-126236/14.
XX
PT New nucleotide sequences isolated from Pyrococcus abyssi encode
PT proteins useful in industry -
XX
PS Claim 7; Pages 1357-1358; 1657pp; French.
XX
CC The present invention relates to the genomic sequence of Pyrococcus
CC abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such P. abyssi protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO2000065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAH75903-AAH75920 and AAG66436.
XX
SQ Sequence 363 AA;
Query Match 1.3%; Score 7; DB 22; Length 363;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 405 FERALSL 411
Db 16 FERALSL 22
|||||

RESULT 52
AAB32774
ID AAB32774 standard; Protein; 370 AA.
XX

```

```

AC AAB32774;
XX
DT 25-JAN-2001 (first entry)
XX
DE Eucalyptus grandis transcription factor protein sequence #232.
XX
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB.
XX
OS Eucalyptus grandis.
XX
PN WO200053724-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US06112.
XX
PR 11-MAR-1999; 99US-0266513.
XX
PR 18-AUG-1999; 99US-0149485.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Wood M, McGrath A, Shenk MA, Glenn M;
XX
WPI; 2000-579369/54.
XX
DR New isolated polynucleotide encoding a plant transcription factor for
XX producing a plant e.g. a woody plant, preferably eucalyptus or pine,
XX having modified gene expression or modified activity of a polypeptide
XX
XX Claim 8; Pages 298-299; 747pp; English.
XX
CC The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
CC transcription factor. The transcription factor may be used to produce a
CC plant having modified gene expression such as a woody plant e.g. a
CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
CC to modify the activity of a polypeptide in a plant. The transcription
CC factors of the present invention are members from the following families
CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
CC zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2
CC Cys2His2, CCAAT box elements and MYB.
XX
SQ Sequence 370 AA;
Query Match 1.3%; Score 7; DB 21; Length 370;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 409 LSLAENE 415
Db 323 LSLAENE 329
|||||

RESULT 53
AAG48637
ID AAG48637 standard; Protein; 373 AA.
XX
AC AAG48637;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 61441.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

```

XX Arabidopsis thaliana.
OS EP1033405-A2.
XX PD 06-SEP-2000.
PN PD 25-FEB-2000; 2000EP-0301439.
XX PF 25-FEB-1999; 99US-0121825.
XX PF 05-MAR-1999; 99US-0123180.
XX PF 09-MAR-1999; 99US-0123548.
XX PF 23-MAR-1999; 99US-0125788.
XX PF 25-MAR-1999; 99US-0126264.
XX PF 29-MAR-1999; 99US-0126785.
XX PF 01-APR-1999; 99US-0127462.
XX PF 06-APR-1999; 99US-0128234.
XX PF 08-APR-1999; 99US-0128714.
XX PF 16-APR-1999; 99US-0129845.
XX PF 19-APR-1999; 99US-0130077.
XX PF 21-APR-1999; 99US-0130449.
XX PF 23-APR-1999; 99US-0130510.
XX PF 28-APR-1999; 99US-0130891.
XX PF 30-APR-1999; 99US-0131449.
XX PF 30-APR-1999; 99US-0132048.
XX PF 30-APR-1999; 99US-0132407.
XX PF 04-MAY-1999; 99US-0132484.
XX PF 05-MAY-1999; 99US-0132485.
XX PF 06-MAY-1999; 99US-0132486.
XX PF 07-MAY-1999; 99US-0132487.
XX PF 11-MAY-1999; 99US-0132456.
XX PF 14-MAY-1999; 99US-0132418.
XX PF 14-MAY-1999; 99US-0132419.
XX PF 14-MAY-1999; 99US-0132421.
XX PF 14-MAY-1999; 99US-0134370.
XX PF 18-MAY-1999; 99US-0134768.
XX PF 19-MAY-1999; 99US-0134941.
XX PF 20-MAY-1999; 99US-0135124.
XX PF 21-MAY-1999; 99US-0135353.
XX PF 24-MAY-1999; 99US-0135629.
XX PF 25-MAY-1999; 99US-0136021.
XX PF 27-MAY-1999; 99US-0136392.
XX PF 28-MAY-1999; 99US-0136782.
XX PF 01-JUN-1999; 99US-0137222.
XX PF 03-JUN-1999; 99US-0137528.
XX PF 04-JUN-1999; 99US-0137502.
XX PF 07-JUN-1999; 99US-0137724.
XX PF 08-JUN-1999; 99US-0138094.
XX PF 10-JUN-1999; 99US-0138540.
XX PF 10-JUN-1999; 99US-0138847.
XX PF 14-JUN-1999; 99US-0139119.
XX PF 16-JUN-1999; 99US-0139452.
XX PF 16-JUN-1999; 99US-0139453.
XX PF 17-JUN-1999; 99US-0139492.
XX PF 18-JUN-1999; 99US-0139454.
XX PF 18-JUN-1999; 99US-0139455.
XX PF 18-JUN-1999; 99US-0139456.
XX PF 18-JUN-1999; 99US-0139457.
XX PF 18-JUN-1999; 99US-0139458.
XX PF 18-JUN-1999; 99US-0139459.
XX PF 18-JUN-1999; 99US-0139460.
XX PF 18-JUN-1999; 99US-0139461.
XX PF 18-JUN-1999; 99US-0139462.
XX PF 18-JUN-1999; 99US-0139463.
XX PF 18-JUN-1999; 99US-0139750.
XX PF 18-JUN-1999; 99US-0139763.
XX PF 21-JUN-1999; 99US-0139817.
XX PF 22-JUN-1999; 99US-0139899.
XX PF 23-JUN-1999; 99US-0140353.
XX PF 23-JUN-1999; 99US-0140354.
XX PF 24-JUN-1999; 99US-0140695.
XX PF 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.

```
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 7; DB 21; Length 373;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 ENEEEAA 419
DB 341 ENEEEAA 347
|||||

RESULT 54
ABB47970
ID ABB47970 standard; Protein; 373 AA.
XX
AC ABB47970;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #674.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR01118.
```

```
XX 11-APR-2000; 2000FR-0004629.
XX (INSP ) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX WPI; 2002-010914/01.
XX
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides
XX
PS Claim 6; SEQ ID No 675; 192pp; French.
XX
CC The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 373 AA;

Query Match 1.3%; Score 7; DB 23; Length 373;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 443 RLALVNN 449
DB 152 RLALVNN 158
|||||

RESULT 55
AAG25383
ID AAG25383 standard; Protein; 383 AA.
XX
AC AAG25383;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 29425.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
```

XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139751.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 7; DB 21; Length 383;
Best Local Similarity 100.08; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 413 ENEEEAA 419
|||||
Db 351 ENEEEAA 357

RESULT 56
AAG48636
ID AAG48636 standard; Protein; 384 AA.

AC AAG48636;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 61440.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

XX Claim 20; SEQ ID No 45017; 103pp; English.

PS The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 387 AA;

Query Match 1.3%; Score 7; DB 22; Length 387;

Best Local Similarity 100.0%; Pred. No. 4.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 DRPLSW 21
|||||

Db 346 DRPLSW 352

RESULT 58

AAAG48625

ID AAG48625 standard; Protein; 392 AA.

XX AC AAG48625;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 61425.

DE Arabidopsis thaliana protein fragment SEQ ID NO: 61425.

XX Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX Arabidopsis thaliana.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

```
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 7; DB 21; Length 392;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 ENEEEAA 419
Db 360 ENEEEAA 366
|||||

RESULT 59
AAG25382
ID AAG25382 standard; Protein; 394 AA.
XX
AC AAG25382;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 29424.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
```


PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135623.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137508.
PR 04-JUN-1999; 99US-0137508.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 7; DB 21; Length 394;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 ENEEEAA 419
Db 362 ENEEEAA 368
|||||

RESULT 60
ABB63456
ID ABB63456 standard; Protein; 398 AA.
XX
AC ABB63456;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 17160.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL07559.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 17160; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 398 AA;

Query Match 1.3%; Score 7; DB 21; Length 394;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 ENEEEAA 419
Db 362 ENEEEAA 368
|||||

RESULT 60
ABB63456
ID ABB63456 standard; Protein; 398 AA.
XX
AC ABB63456;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 17160.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL07559.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 17160; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 398 AA;

Query Match 1.3%; Score 7; DB 22; Length 398;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 524 QHLIKQL 530
Db 20 QHLIKQL 26
|||||

RESULT 61
ABB62489
ID ABB62489 standard; Protein; 401 AA.
XX
AC ABB62489;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 14259.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL06592.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 14259; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 401 AA;

Query Match 1.3%; Score 7; DB 22; Length 401;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 337 YKEVLKQ 343
Db 292 YKEVLKQ 298
|||||

RESULT 62
AAG48624
ID AAG48624 standard; Protein; 403 AA.
XX

AC AAG48624;
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 61424.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 200EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0132566.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0158293.
 PR 13-OCT-1999; 99US-0158294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 26-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 7; DB 21; Length 403;
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 ENEEEAA 419
 |||||
 Db 371 ENEEEAA 377

RESULT 63
 AAB33303
 ID AAB33303 standard; Protein; 412 AA.
 XX
 AC AAB33303;
 XX

DT 25-JAN-2001 (first entry)
 XX
 DE Pinus radiata transcription factor protein sequence #343.
 XX
 PF Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
 KW

KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
 KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
 KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
 KW type 2 Cys2His2; CCAAT box element; MYB.
 XX
 OS Pinus radiata.
 PN WO2000053724-A2.
 XX
 PD 14-SEP-2000.
 XX
 XX 09-MAR-2000; 2000WO-US06112.
 PF
 XX 11-MAR-1999; 99US-0266513.
 PR 18-AUG-1999; 99US-0149485.
 XX
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 PI Wood M, McGrath A, Shenk MA, Glenn M,
 XX WPI; 2000-579369/54.
 DR
 XX New isolated polynucleotide encoding a plant transcription factor for
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
 PT having modified gene expression or modified activity of a polypeptide
 PT
 XX Claim 8; Pages 715-716; 747pp; English.
 PS
 CC The present invention relates to novel plant transcription factors from
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
 CC transcription factor. The transcription factor may be used to produce a
 CC plant having modified gene expression such as a woody plant e.g. a
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
 CC to modify the activity of a polypeptide in a plant. The transcription
 CC factors of the present invention are members from the following families
 CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
 CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
 CC zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2
 CC Cys2His2, CCAAT box elements and MYB.
 XX
 SQ Sequence 412 AA;
 Query Match 1.3%; Score 7; DB 21; Length 412;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 373 RLLQWGI 379
 |||||
 Db 116 RLLQWGI 122
 RESULT 64
 ABP78893
 ID ABP78893 standard; Protein; 414 AA.
 XX
 AC ABP78893;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE N. gonorrhoeae amino acid sequence SEQ ID 4316.
 XX
 KW Antibacterial; infection; vaccine; gene therapy.
 XX
 OS Neisseria gonorrhoeae.
 XX
 PN WO200279243-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 12-FEB-2002; 2002WO-IB02069.
 XX

```
PR 12-FEB-2001; 2001GB-0003424.
XX (CHIR-) CHIRON SPA.
XX Fontana MR, Pizza M, Massignani V, Monaci E;
PI WPI; 2003-058415/05.
XX N-PSDB; ABZ39863.
DR New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection -
XX Disclosure; Page 503; 815pp; English.
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention.
XX
SQ Sequence 414 AA;
  Query Match 1.3%; Score 7; DB 24; Length 414;
  Best Local Similarity 100.0%; Pred. No. 4.3e+02;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 84 VYIDEID 90
Db 181 VYIDEID 187
  |||||
  |||||

RESULT 65
ABBI2386
ID ABBI2386 standard; Protein; 419 AA.
XX
AC ABBI2386;
XX
DT 15-JAN-2002 (first entry)
XX
DE Human bone marrow expressed protein SEQ ID NO: 141.
XX
KW Human; bone marrow; cytostatic; antirheumatic; antiarthritic; vulnary;
KW antiinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;
KW antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;
KW antitumor; fungicide; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; cerebroprotective; antianaemic; infection;
KW nervous system disorder; autoimmune disorder; inflammation; allergy.
XX
XX Homo sapiens.
OS
XX
XX WO200174836-A1.
PN
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US10472.
PF
XX
XX 31-MAR-2000; 2000US-0540217.
PR
XX 23-AUG-2000; 2000US-0649167.
PR
XX 23-AUG-2000; 2000US-0649267.
PR
XX 30-NOV-2000; 2000US-250583P.
PR
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT, Ford JE, Boyle BJ;
PI
XX WPI; 2001-626375/72.
DR N-PSDB; ABA09631.
XX
XX New bone marrow-expressed nucleic acids and polypeptides, useful for
PT diagnosis, treatment of inflammatory, autoimmune, neurological, cancer
```

```
PT and increasing hematopoiesis, stem cell survival and bone growth and
PT remodeling -
XX Claim 10; Page 238-239; 380pp; English.
PS
XX The present invention relates to bone marrow expressed polynucleotides
CC and proteins. These sequences can be used in the treatment of
CC inflammatory conditions (eg arthritis, Crohn's disease), cancer, central
CC and peripheral nervous system diseases and neuropathies, such as
CC Alzheimer's, Parkinson's and Huntington's diseases, spinal cord
CC disorders, head trauma, cerebrovascular diseases, myeloid and lymphoid
CC cell disorders, platelet disorders, stem cell disorders, bone
CC degenerative disorders, autoimmune disorders, for example multiple
CC sclerosis, diabetes and arthritis, viral and bacterial infections,
CC allergies and blood coagulation disorders. The present sequence is a
CC protein of the invention.
XX
SQ Sequence 419 AA;
  Query Match 1.3%; Score 7; DB 22; Length 419;
  Best Local Similarity 100.0%; Pred. No. 4.4e+02;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 76 KARALTE 82
Db 369 KARALTE 375
  |||||
  |||||

RESULT 66
AAG48635
ID AAG48635 standard; Protein; 420 AA.
XX
AC AAG48635;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 61439.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR
XX 05-MAR-1999; 99US-0123180.
PR
XX 09-MAR-1999; 99US-0123548.
PR
XX 23-MAR-1999; 99US-0125788.
PR
XX 25-MAR-1999; 99US-0126264.
PR
XX 29-MAR-1999; 99US-0126785.
PR
XX 01-APR-1999; 99US-0127462.
PR
XX 06-APR-1999; 99US-0128234.
PR
XX 08-APR-1999; 99US-0128714.
PR
XX 16-APR-1999; 99US-0129845.
PR
XX 19-APR-1999; 99US-0130077.
PR
XX 21-APR-1999; 99US-0130449.
PR
XX 23-APR-1999; 99US-0130510.
PR
XX 28-APR-1999; 99US-0131449.
PR
XX 30-APR-1999; 99US-0132048.
PR
XX 30-APR-1999; 99US-0132407.
PR
XX 04-MAY-1999; 99US-0132484.
PR
XX 05-MAY-1999; 99US-0132485.
PR
XX 06-MAY-1999; 99US-0132486.
PR
XX 07-MAY-1999; 99US-0132487.
PR
XX 11-MAY-1999; 99US-0132863.
PR
XX 99US-0134256.
```

PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145919.
PR	18-MAY-1999;	99US-0134768.	PR	28-JUL-1999;	99US-0145951.
PR	19-MAY-1999;	99US-0134941.	PR	02-AUG-1999;	99US-0146386.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146388.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146389.
PR	24-MAY-1999;	99US-0135629.	PR	03-AUG-1999;	99US-0147038.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147204.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147302.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	12-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149923.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-01

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 1.3%; Score 7; DB 21; Length 420;

Best Local Similarity 100.0%; Pred.No. 4.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 413 ENEEEAA 419
Db 388 ENEEEAA 394
|||||

RESULT 67

AAG25381
ID AAG25381 standard; Protein; 430 AA.

XX AAG25381;

XX AAG25381;

XX AAG25381;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

DE Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29423.

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.

```
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 7; DB 21; Length 430;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 ENEEEA 419
Db 398 ENEEEA 404
|||||

RESULT 68
AAU48807
ID AAU48807 standard; Protein; 433 AA.
XX
AC AAU48807;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #9703.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
WPI; 2001-616774/71.
N-PSDB; AAS59543.

Propionibacterium acnes polypeptides and nucleic acids useful for
vaccinating against and diagnosing infections, especially useful for
treating acne vulgaris -

Example 1; SEQ ID No 10002; 1069pp; English.

Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
polypeptides. The proteins and their associated DNA sequences are used in
the treatment, prevention and diagnosis of medical conditions caused by
P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
P. acnes is also involved in infections of bone, joints and the central
nervous system, however it is particularly involved in the inflammatory
lesions associated with acne vulgaris. A method for detecting the
presence or absence of P. acnes in a patient comprises contacting a
sample with a binding agent that binds to the proteins of the invention
and determining the amount of bound protein in the sample. The
polypeptides may be used as antigens in the production of antibodies
specific for P. acnes proteins. These antibodies can be used to
downregulate expression and activity of P. acnes polypeptides and
therefore treat P. acnes infections. The antibodies may also be used as
diagnostic agents for determining P. acnes presence, for example, by
enzyme linked immunosorbent assay (ELISA).

Note: The sequence data for this patent did not form part of the printed
```


CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 433 AA;

Query Match 1.3%; Score 7; DB 22; Length 433;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 VYIDEID 90

|||||

Db 191 VYIDEID 197

RESULT 69

AAW41733

ID AAW41733 standard; Protein; 437 AA.

XX AC AAW41733;

XX DT 06-JUL-1998 (first entry)

XX DE Arabidopsis chloroplast division protein AtFtsZ.

XX KW Chloroplast division; plastid division; FtsZ protein; AtFtsZ;
transgenic plant.

XX OS Arabidopsis thaliana.

XX PN WO9800436-A1.

XX PD 08-JAN-1998.

XX PF 27-JUN-1997; 97WO-US11287.

XX PR 28-JUN-1996; 96US-0020959.

XX PA (UYNE-) UNIV NEVADA.

XX PI Osteryoung KW;

XX WPI; 1998-086896/08.

XX DR N-PSDB; AAV05550.

XX Arabidopsis thaliana plant plastid division genes designated cpFtsZ
and AtFtsZ - useful for obtaining transgenic plants with novel
phenotype(s) characterised by alterations in plastid number and size

XX Example 1; Page 24-25; 39pp; English.

CC This Arabidopsis AtFtsZ protein is the putative product of an
isolated AtFtsZ gene (see AAV05550) identified on the basis of
homology to bacterial FtsZ (filamenting temperature sensitive)
sequences and homology to cpFtsZ (see AAW41732), another FtsZ
homologue of Arabidopsis. The AtFtsZ protein contains a
glycine-rich 'tubulin signature' motif which is conserved among
FtsZ proteins and tubulins and which is important for GTP binding.
CC This suggests that AtFtsZ may have a function analogous to the
cytoskeletal role of tubulin, which requires GTP-dependent
polymerisation for its activity. The AtFtsZ protein may not be
full-length; it lacks the N-terminal transit peptide sequence of
cpFtsZ. The claimed AtFtsZ and cpFtsZ genes (see also AAV05549) are
useful for obtaining transgenic plants with novel phenotypes
characterised by alterations in plastid number and size.

SQ Sequence 437 AA;

Query Match 1.3%; Score 7; DB 19; Length 437;

Best Local Similarity 100.0%; Pred. No. 4.6e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 508 VAAQKSE 514

|||||

Db 7 VAAQKSE 13

RESULT 70

AAG48623

ID AAG48623 standard; Protein; 439 AA.

XX AC AAG48623;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 61423.

XX KW Hybridisation assay; signal transduction pathway; metabolic pathway;
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 06-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 14-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134768.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

XX 24-MAY-1999; 99US-0135629.

XX 25-MAY-1999; 99US-0136021.

XX 27-MAY-1999; 99US-0136392.

XX 28-MAY-1999; 99US-0136782.

XX 01-JUN-1999; 99US-0137222.

XX 03-JUN-1999; 99US-0137528.

XX 04-JUN-1999; 99US-0137502.

XX 07-JUN-1999; 99US-0137724.

XX 08-JUN-1999; 99US-0138094.

XX 10-JUN-1999; 99US-0138540.

XX 10-JUN-1999; 99US-0138847.

XX 14-JUN-1999; 99US-0139119.

XX 16-JUN-1999; 99US-0139452.

XX 17-JUN-1999; 99US-0139453.

XX 18-JUN-1999; 99US-0139454.

XX 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.3%; Score 7; DB 21; Length 439;

Best Local Similarity 100.0%; Pred. No. 4.6e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 ENEEEAA 419
| | | | |
Db 407 ENEEEAA 413

RESULT 71
ABP80005
ID ABP80005 standard; Protein; 470 AA.
XX

AC ABP80005;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae amino acid sequence SEQ ID 6540.
XX
KW Antibacterial; infection; vaccine; gene therapy.
XX
KW Neisseria gonorrhoeae.
OS
XX WO200279243-A2.
XX
XX 10-OCT-2002.
XX
XX 12-FEB-2002; 2002WO-IB02069.
XX
PR 12-FEB-2001; 2001GB-0003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Pizzo M, Massignani V, Monaci E;
XX
DR WPI; 2003-058415/05.
DR N-PSDB; ABZ40975.
XX
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection -
PT
XX Disclosure; Page 663-664; 815pp; English.
XX
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention.
XX
XX
SQ Sequence 470 AA;

Query Match 1.3%; Score 7; DB 24; Length 470;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 YIDEIDV 91
Db 302 YIDEIDV 308
|||||

RESULT 72
AAW00290
ID AAW00290 standard; Protein; 507 AA.
XX
AC AAW00290;
XX
DT 20-NOV-1996 (first entry)
XX
DE Mature Penicillin V amidohydrolase.
XX
KW Penicillin V amidohydrolase; PVA; F. oxysporum; strain 435;
KW hydrolysis; phenoxy-methylpenicillin; 6-aminopenicillanic acid;
KW 6-APA; beta-lactam; semi-synthetic penicillin; expression vector;
KW recombinant production.
XX
OS Fusarium oxysporum.
XX
XX US5516679-A.
PN
XX 14-MAY-1996.
XX
XX 23-DEC-1994; 94US-0363475.
PF
XX

PR 23-DEC-1994; 94US-0363475.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Burnett WV, Chiang S, Tonzi SM;
XX
XX WPI; 1996-251011/25.
DR N-PSDB; AAT40245.
XX
XX DNA encoding penicillin V amidohydrolase (PVA) from Fusarium
PT oxysporum - also recombinant vectors and host cells for production
PT of PVA for use in the manufacture of penicillin.
XX
XX Claim 1; Column 31-34; 46pp; English.
XX
XX The sequences given in AAW00290-91 represent the mature and full length
CC penicillin V amidohydrolase (PVA) sequences from F. oxysporum strain
CC 435 respectively. PVA is used for the enzymatic hydrolysis of
CC penicillin V (phenoxy-methylpenicillin) to 6-aminopenicillanic acid
CC (6-APA). 6-APA is the active beta-lactam nucleus used in the
CC manufacture of semi-synthetic penicillins. The PVA coding sequences can
CC be inserted into expression vectors for the recombinant production of
CC PVA in a suitable host, pref. Fusarium sp.
XX
XX
SQ Sequence 507 AA;

Query Match 1.3%; Score 7; DB 17; Length 507;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 115 SLKLPGT 121
Db 15 SLKLPGT 21
|||||

RESULT 73
AAW00291
ID AAW00291 standard; Protein; 532 AA.
XX
AC AAW00291;
XX
XX 20-NOV-1996 (first entry)
DT
XX Full length Penicillin V amidohydrolase.
DE
XX Penicillin V amidohydrolase; PVA; F. oxysporum; strain 435;
KW hydrolysis; phenoxy-methylpenicillin; 6-aminopenicillanic acid;
KW 6-APA; beta-lactam; semi-synthetic penicillin; expression vector;
KW recombinant production.
XX
OS Fusarium oxysporum.
XX
XX
FH Key Location/Qualifiers
FT Peptide 1..25
FT Protein /note= "Signal peptide"
FT /note= "Mature PVA"
XX
XX US5516679-A.
PN
XX 14-MAY-1996.
XX
XX 23-DEC-1994; 94US-0363475.
PF
XX 23-DEC-1994; 94US-0363475.
PR
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Burnett WV, Chiang S, Tonzi SM;
XX
XX WPI; 1996-251011/25.
DR N-PSDB; AAT40245.
XX

PT DNA encoding penicillin V amidohydrolase (PVA) from *Fusarium*
PT *oxy-sporum* - also recombinant vectors and host cells for production
PT of PVA for use in the manufacture of penicillin.
XX
XX
PS Claim 1; Column 37-42; 46pp; English.
XX
XX The sequences given in AAW00290-91 represent the mature and full length
CC penicillin V amidohydrolase (PVA) sequences from *F. oxysporum* strain
CC 435 respectively. PVA is used for the enzymatic hydrolysis of
CC penicillin V (phenoxymethylpenicillin) to 6-aminopenicillanic acid
CC (6-APA). 6-APA is the active beta-lactam nucleus used in the
CC manufacture of semi-synthetic penicillins. The PVA coding sequences can
CC be inserted into expression vectors for the recombinant production of
CC PVA in a suitable host, pref. *Fusarium* sp.
XX
XX
SQ Sequence 532 AA;
Query Match 1.3%; Score 7; DB 17; Length 532;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 115 SLKLPQT 121
Db 40 SLKLPQT 46
|||||
RESULT 74
AA34962
ID AA34962 standard; Protein; 544 AA.
XX
XX AC AA34962;
XX
DT 13-SEP-1999 (first entry)
XX
DE Chlamydia pneumoniae surface exposed polypeptide.
XX
KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX
OS Chlamydia pneumoniae.
XX
XX WO927105-A2.
XX
PD 03-JUN-1999.
XX
XX 20-NOV-1998; 98WO-IB01890.
XX
XX 04-NOV-1998; 98US-0107078.
XX
XX 21-NOV-1997; 97FR-0014673.
XX
XX (GIST) GENSET.
XX
XX Griffais R;
XX
XX WPI; 1999-357842/30.
XX
XX Genome sequence of *Chlamydia pneumoniae*
XX
XX Page 880-881; Disclosure; 1912pp; English.
XX
XX AA34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AA34584-Y35879) of *Chlamydia pneumoniae*.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see AA34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX

SQ Sequence 544 AA;
Query Match 1.3%; Score 7; DB 20; Length 544;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 528 KQLRQHF 534
Db 225 KQLRQHF 231
|||||
RESULT 75
ABJ18511
ID ABJ18511 standard; Protein; 565 AA.
XX
XX AC ABJ18511;
XX
DT 13-FEB-2003 (first entry)
XX
DE Human Cryptovirus strain BBR protein region SEQ ID NO 22.
XX
KW Neuroprotective; virucide; anticonvulsant; antiinflammatory; Cryptovirus;
KW idiopathic; neurological; neurodegenerative; neuropsychological; vaccine;
KW epilepsy; neuropsychiatric disorder; subacute sclerosing panencephalitis;
KW epileptiform disease; multiple sclerosis; chronic fatigue syndrome;
KW primary lymphadenopathy-associated illness; gene therapy.
XX
XX Rubulavirus sp.
OS
XX WO200277211-A2.
XX
XX 03-OCT-2002.
XX
XX 07-FEB-2002; 2002WO-US04117.
XX
XX 07-FEB-2001; 2001US-267253P.
XX
XX (CRYP-) CRYPTIC AFFLICTIONS LLC.
XX
XX Robbins SJ;
XX
XX WPI; 2003-040586/03.
XX
XX N-PSDB; APT13852.
XX
XX New nucleic acid useful for diagnosing and treating idiopathic
PT neurological disorders, including epileptiform diseases, e.g. epilepsy,
PT and lymphadenopathy-associated illnesses, and in screening of potential
PT new antiviral drugs -
XX
XX Disclosure; Page 54-55; 262pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising: contiguous
CC nucleotide positions 1-15246 of a sequence comprising 15246 bp fully
CC defined in the specification; a nucleotide sequence complementary to the
CC 15246bp sequence; or a Cryptovirus-specific fragment of the 15246bp or
CC its complement, at least about 5 nucleotides long. The nucleic acid is
CC useful in diagnosing and treating many idiopathic neurological,
CC neurodegenerative, neuropsychological and neuropsychiatric disorders,
CC including epileptiform diseases (e.g. epilepsy, multiple sclerosis, and
CC chronic fatigue syndrome or subacute sclerosing panencephalitis) and
CC primary lymphadenopathy-associated illnesses, and in research and
CC development, including screening of potential new antiviral drugs. The
CC nucleic acid, protein and the viral particle are useful in manufacturing
CC a vaccine. The protein is also used in producing a Cryptovirus-specific
CC antibody. The antibody may also be used in manufacturing a medicament for
CC the treatment of Cryptovirus infections. The polynucleotides of the
CC invention can be used to treat disorders by gene therapy. This sequence
CC represents a human Cryptovirus strain BBR protein region of the
CC invention.
XX
XX Sequence 565 AA;
SQ
Query Match 1.3%; Score 7; DB 24; Length 565;

Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 108 QVPRGT 114
Db 439 QVPRGT 445
|||||

Search completed: February 10, 2004, 13:27:40
Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 13:26:36 ; Search time 20 Seconds
(without alignments)
2582.129 Million cell updates/sec

Title: US-09-821-812-5

Perfect score: 537

Sequence: 1 GGLTTTIGRLGVRPRLS.....PDHVDTHLKLRLQHFAML 537

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	1.5	293	1 B9AG55	virB9 protein prec
2	8	1.5	532	2 D71267	probable UDP-N-ace
3	8	1.5	1332	1 I48314	homeotic protein C
4	7	1.3	59	2 A51363	hypothetical prote
5	7	1.3	70	2 S56769	RNA-directed RNA p
6	7	1.3	75	1 R3EC18	ribosomal protein
7	7	1.3	75	2 E84076	ribosomal protein
8	7	1.3	75	2 A80430	30S ribosomal prot
9	7	1.3	75	2 E84995	30S ribosomal subu
10	7	1.3	75	2 B86117	30S ribosomal subu
11	7	1.3	75	2 B91276	30S ribosomal subu
12	7	1.3	75	2 A11052	30S ribosomal chai
13	7	1.3	92	2 T49763	probable magnapori
14	7	1.3	110	2 S11786	nolE protein - Rhi
15	7	1.3	116	2 B75456	hypothetical prote
16	7	1.3	116	2 C72654	hypothetical prote
17	7	1.3	133	2 D70105	hypothetical prote
18	7	1.3	140	2 S75420	ribosomal protein
19	7	1.3	147	2 C70923	hypothetical prote
20	7	1.3	147	2 G72640	hypothetical prote
21	7	1.3	151	2 G84131	hypothetical prote
22	7	1.3	174	2 S73113	hypothetical prote
23	7	1.3	183	2 T16186	hypothetical prote
24	7	1.3	209	2 E83270	hypothetical prote
25	7	1.3	220	2 C64146	hypothetical prote
26	7	1.3	221	2 JE0171	ribonuclease T2 (B
27	7	1.3	223	2 T37032	probable integral
28	7	1.3	229	2 I40068	probable ABC-type
29	7	1.3	234	2 AF0504	conserved hypothet

30	7	1.3	245	2 C84017	hypothetical prote
31	7	1.3	262	2 H75377	conserved hypothet
32	7	1.3	264	2 S28660	proliferating cell
33	7	1.3	267	2 AD1835	hypothetical prote
34	7	1.3	272	2 QJ1802	B8R 31K protein pr
35	7	1.3	272	2 G42526	probable formamido
36	7	1.3	274	2 T51714	hypothetical prote
37	7	1.3	281	2 S69525	hypothetical prote
38	7	1.3	282	2 D72679	hypothetical cytos
39	7	1.3	283	2 AH3387	probable transcrip
40	7	1.3	284	2 C83237	fimbrial adhesin K
41	7	1.3	285	1 YQEC88	hypothetical prote
42	7	1.3	287	2 B87045	ribonuclease (EC 3
43	7	1.3	288	2 H70017	tissue factor prec
44	7	1.3	292	1 KFRB3	probable transport
45	7	1.3	296	2 T45203	aromatic metabolit
46	7	1.3	303	1 B69903	type II site-speci
47	7	1.3	315	2 S72472	type II site-speci
48	7	1.3	315	2 AH2259	homoserine kinase
49	7	1.3	316	2 H82958	probable cobalamin
50	7	1.3	317	2 B82084	homoserine O-acety
51	7	1.3	319	2 D69110	hypothetical prote
52	7	1.3	331	2 G87494	hypothetical prote
53	7	1.3	331	2 D96515	transcription regu
54	7	1.3	337	2 AD2837	hypothetical prote
55	7	1.3	337	2 F97614	hypothetical prote
56	7	1.3	341	1 LUPY1	annexin I type 1 -
57	7	1.3	343	1 A44118	annexin I type 2 -
58	7	1.3	344	2 G82690	aspartate-B-semial
59	7	1.3	348	2 F71372	probable protoporp
60	7	1.3	358	2 A98332	malate dehydrogena
61	7	1.3	359	2 S17905	hypothetical prote
62	7	1.3	361	2 D86327	hypothetical prote
63	7	1.3	363	2 B75132	cell division cont
64	7	1.3	365	2 C87684	conserved hypothet
65	7	1.3	367	2 E69072	hypothetical prote
66	7	1.3	369	2 A71025	hypothetical prote
67	7	1.3	372	2 C81263	probable integral
68	7	1.3	373	2 AD1256	conserved hypothet
69	7	1.3	373	2 AH1618	conserved hypothet
70	7	1.3	374	2 S48639	fructose-bisphosph
71	7	1.3	389	2 AG2460	hypothetical prote
72	7	1.3	390	2 E96565	hypothetical prote
73	7	1.3	390	2 T51713	probable formamido
74	7	1.3	399	2 T35440	probable polyamine
75	7	1.3	412	2 A70416	ATP-dependent clip
76	7	1.3	412	2 F87460	hypothetical prote
77	7	1.3	414	2 D81851	ATP-dependent clip
78	7	1.3	414	2 A81091	ATP-dependent clip
79	7	1.3	419	2 C82109	Glutamyl-tRNA redu
80	7	1.3	420	2 D87492	hypothetical prote
81	7	1.3	423	2 AD0383	ATP-dependent clip
82	7	1.3	423	2 AD0558	ATP-dependent clip
83	7	1.3	424	2 D90690	ATP-dependent clip
84	7	1.3	424	2 H85540	ATP-dependent clip
85	7	1.3	424	2 A48709	ATP-dependent clip
86	7	1.3	426	2 F83420	ATP-dependent clip
87	7	1.3	426	2 B82712	ATP-dependent clip
88	7	1.3	426	2 F82139	ATP-dependent clip
89	7	1.3	428	2 S75076	hypothetical prote
90	7	1.3	429	2 E84985	hypothetical prote
91	7	1.3	436	2 A83909	hypothetical prote
92	7	1.3	437	2 T21518	hyurenine-oxogluc
93	7	1.3	437	2 A71701	osmolarity sensor
94	7	1.3	445	2 AB2022	hypothetical prote
95	7	1.3	447	2 C83683	phosphoglucosamin
96	7	1.3	451	2 E70013	hypothetical prote
97	7	1.3	458	2 T29520	hypothetical prote
98	7	1.3	462	2 D87630	conserved hypothet
99	7	1.3	470	2 D81063	Trk system potassi
100	7	1.3	478	2 JC7770	chloroplast divisi

ALIGNMENTS

RESULT 1

B9AG55
 virB9 protein precursor - Agrobacterium tumefaciens plasmids pTi15955 and pTiA6
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
 C:Accession: S00785; B30402; A27647; A35737; A26217
 R:Thompson, D.V.; Melcher, L.S.; Idler, K.B.; Schilperoort, R.A.; Hooykaas, P.J.J.
 Nucleic Acids Res. 16, 4621-4636, 1988
 A:Title: Analysis of the complete nucleotide sequence of the Agrobacterium tumefaciens virB9 gene
 A:Reference number: S00777; MUID:88247765; PMID:2837739
 A:Accession: S00785
 A:Molecule type: DNA
 A:Residues: 1-293 <DNA>
 A:Cross-references: EMBL:X06826; NID:g93195; PIDN:CAA29979.1; PID:g757730
 A:Experimental source: strain 15955, plasmid pTi15955
 R:Ward, J.E.; Akiyoshi, D.E.; Regier, D.; Datta, A.; Gordon, M.P.; Nester, E.W.
 submitted to GenBank, December 1989
 A:Reference number: A30402
 A:Accession: B30402
 A:Molecule type: DNA
 A:Residues: 1-293 <WAR>
 A:Cross-references: GB:J03216; NID:q1196971; PIDN:AAA88654.1; PID:g1196981
 A:Experimental source: plasmid pTiA6
 R:Ward, J.E.; Akiyoshi, D.E.; Regier, D.; Datta, A.; Gordon, M.P.; Nester, E.W.
 J. Biol. Chem. 263, 5804-5814, 1988
 A:Title: Characterization of the virB operon from an Agrobacterium tumefaciens Ti plasmid
 A:Reference number: A28621; MUID:88186901; PMID:3281947
 A:Accession: A27647
 A:Molecule type: DNA
 A:Residues: 1-271 <WA2>
 A:Cross-references: GB:J03216
 A:Experimental source: plasmid pTiA6
 A:Note: this sequence was designated the amino-terminal portion of ORF 10 in this reference
 R:Ward, J.E.; Akiyoshi, D.E.; Regier, D.; Datta, A.; Gordon, M.P.; Nester, E.W.
 J. Biol. Chem. 265, 4768, 1990
 A:Reference number: A35737; MUID:90170994; PMID:2307685
 A:Contents: erratum
 A:Accession: A35737
 A:Molecule type: DNA
 A:Residues: 97-293 <WA3>
 A:Experimental source: plasmid pTiA6
 C:Genetics:
 A:Genome: plasmid
 C:Superfamily: tumor-inducing plasmid pTiCS8 virB9 protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-293/Product: virB9 protein #status predicted <MAT>

Query Match 1.5%; Score 8; DB 1; Length 293;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 LDQPVTL 300
 |||||
 Db 173 LDQPVTL 180

RESULT 2

D71267
 probable UDP-N-acetyluracilalanine-D-glutamate ligase (murD) - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
 C:Accession: D71267
 R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A:Reference number: A71250; MUID:98332770; PMID:9665876
 A:Accession: D71267
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-532 <COL>
 A:Cross-references: GB:AE001259; GB:AE000520; NID:g3323209; PIDN:AA65856.1; PID:g332321
 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0903

Query Match 1.5%; Score 8; DB 2; Length 532;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 470 EQARALLQ 477
 |||||
 Db 2 EQARALLQ 9

RESULT 3

I48314
 homeotic protein CDP - mouse
 N:Alternate names: CCAAT displacement protein; homeotic protein Cux
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: I48314
 R:Valarche, I.; Tissier-Seta, J.P.; Hirsch, M.R.; Martinez, S.; Goridis, C.; Brunet, J.F.
 Development 119, 881-896, 1993
 A:Title: The mouse homeodomain protein Phox2 regulates Ncam promoter activity in concert
 A:Reference number: I48314; MUID:94244481; PMID:7910552
 A:Accession: I48314
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1332 <RES>
 A:Cross-references: EMBL:X75013; NID:g402589; PIDN:CAA52922.1; PID:g402590
 C:Genetics:
 A:Gene: Cux
 C:Superfamily: homeotic protein CDP; cut repeat homology; homeobox homology
 C:Keywords: DNA binding; duplication; homeobox; nucleus; transcription regulation
 F:365-438/Domain: cut repeat homology <CU1>
 F:755-827/Domain: cut repeat homology <CU2>
 F:938-1010/Domain: cut repeat homology <CU3>
 F:1057-1113/Domain: homeobox homology <HOM>

Query Match 1.5%; Score 8; DB 1; Length 1332;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 PSLAAPAA 32
 |||||
 Db 1230 PSLAAPAA 1237

RESULT 4

AE1363
 hypothetical protein lmo2309 [imported] - Listeria monocytogenes (strain EGD-e)
 C:Species: Listeria monocytogenes
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AE1363
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
 A:Title: Comparative genomics of Listeria species.
 A:Reference number: AE1077; MUID:21537279; PMID:11679669
 A:Accession: AE1363
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-59 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAD00387.1; PID:gl6411779; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo2309

Query Match 1.3%; Score 7; DB 2; Length 59;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 412 AENEESA 418
 Db 17 AENEESA 23
 |||||

RESULT 5
 S56769
 RNA-directed RNA polymerase (EC 2.7.7.48) - Leishmania RNA virus 1-13 (fragment)
 A;Title: Leishmania RNA virus 1-13
 C;Species: Leishmania RNA virus 1-13
 C;Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 18-Jul-1997
 C;Accession: S56769
 R;Widmer, G.; Dooley, S.
 Nucleic Acids Res. 23, 2300-2304, 1995
 A;Title: Phylogenetic analysis of Leishmania RNA virus and Leishmania suggests ancient
 A;Reference number: S56768; UID:95334386; PMID:7610059
 A;Accession: S56769
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-70 <WID>
 A;Cross-references: EMBL:L39069
 C;Keywords: nucleotidyltransferase

Query Match 1.3%; Score 7; DB 2; Length 70;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 EIDVDQE 94
 Db 17 EIDVDQE 23
 |||||

RESULT 6
 R3EC18
 ribosomal protein S18 [validated] - Escherichia coli (strain K-12)
 C;Species: Escherichia coli
 C;Date: 24-Apr-1984 #sequence_revision 24-May-1996 #text_change 01-Mar-2002
 C;Accession: S56747; A02741; E65231
 R;Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.
 Nucleic Acids Res. 23, 2105-2119, 1995
 A;Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.
 A;Reference number: S56314; UID:95334362; PMID:7610040
 A;Accession: S56747
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-75 <BUR>
 A;Cross-references: EMBL:U14003; NID:gl263172; PIDN:AAA97098.1; PID:g537043
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
 R;Yaguchi, M.
 FEBS Lett. 59, 217-220, 1975
 A;Title: Primary structure of protein S18 from the small Escherichia coli ribosomal sub
 A;Reference number: A02741; UID:76210737; PMID:776663
 A;Accession: A02741
 A;Molecule type: protein
 A;Residues: 2-15, 'Q', 17-75 <VAG>
 A;Experimental source: strain K
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; UID:97428617; PMID:9278503
 A;Accession: E65231
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-75 <BLAT>
 A;Cross-references: GB:AE000491; GB:U00096; NID:g2367357; PIDN:AACT7159.1; PID:g1790646;
 A;Experimental source: strain K-12, substrain MG1655
 R;Arnold, R.J.; Reilly, J.P.
 Anal. Biochem. 269, 105-112, 1999
 A;Title: Observation of Escherichia coli ribosomal proteins and their posttranslational

A;Reference number: A59071; UID:99196679; PMID:10094780
 A;Contents: annotation; mass spectrographic analysis
 A;Note: mass spectrographic analysis of post-translational modifications; any acid labil-
 C;Comment: The amino end is acetylated by ribosomal-protein-alanine N-acetyltransferase
 C;Genetics:
 A;Gene: rpeR
 A;Map position: 96 min
 C;Complex: the ribosome is composed of the large (50S) and small (30S) subunit; the larg
 S rRNA and 22 distinct proteins
 C;Complex: small subunit ribosomal proteins: S1 (PIR:R3EC1), S2 (PIR:R3EC2), S3 (PIR:R3EC
 R3EC10), S11 (PIR:R3EC11), S12 (PIR:R3EC12), S13 (PIR:R3EC13), S14 (PIR:R3EC14), S15 (PI
 IR:R3EC21), S22 (PIR:C64901) [validated, UID:99196679]
 C;Function:
 A;Description: may be involved in aminoacyl-transfer RNA binding; located at the decodin
 A;Pathway: protein biosynthesis
 C;Superfamily: Escherichia coli ribosomal protein S18
 C;Keywords: acetylated amino end; protein biosynthesis; ribosome
 F;2-75/Product: ribosomal protein S18 #status experimental <WAT>
 F;2/Modified site: acetylated amino end (Ala) (in mature form) #status experimental

Query Match 1.3%; Score 7; DB 1; Length 75;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 YFRRRKF 52
 Db 4 YFRRRKF 10
 |||||

RESULT 7
 E64076
 ribosomal protein S18 - Haemophilus influenzae (strain Rd KW20)
 C;Species: Haemophilus influenzae
 C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 13-Aug-1999
 C;Accession: E64076
 R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
 Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
 D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A;Reference number: A64000; UID:95350630; PMID:7542800
 A;Accession: E64076
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-75 <TIGR>
 A;Cross-references: GB:U32736; GB:L42023; NID:gl573519; PIDN:AAC22203.1; PID:gl573530; T
 C;Superfamily: Escherichia coli ribosomal
 C;Keywords: protein biosynthesis; ribosome

Query Match 1.3%; Score 7; DB 2; Length 75;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 YFRRRKF 52
 Db 4 YFRRRKF 10
 |||||

RESULT 8
 AB0430
 30S ribosomal protein S18 [imported] - Yersinia pestis (strain CO92)
 C;Species: Yersinia pestis
 C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 17-May-2002
 C;Accession: AB0430
 R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A;Reference number: AB0001; UID:21470413; PMID:11586360
 A;Accession: AB0430
 A;Status: preliminary

```

A:Molecule type: DNA
A:Residues: 1-75 <KUR>
A:Cross-references: GB:ALU590842; PIDN:CAC92766.1; PID:gl5981461; GSPDB:GN00175
C:Genetics:
A:Gene: rpsR
C:Superfamily: Escherichia coli ribosomal protein S18

Query Match          1.3%; Score 7; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 YFRRRK 52
   |||||
Db 4 YFRRRK 10

RESULT 9
B84995
30S ribosomal protein S18 [imported] - Buchnera sp. (strain APS)
C:Species: Buchnera sp.
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C:Accession: B84995
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A:Reference number: AB4930; MUID:20445173; PMID:10993077
A:Accession: B84995
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-75 <STO>
A:Cross-references: GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: rpsR; BU563
C:Superfamily: Escherichia coli ribosomal protein S18

Query Match          1.3%; Score 7; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 YFRRRK 52
   |||||
Db 4 YFRRRK 10

RESULT 10
B86117
30S ribosomal subunit protein S18 [imported] - Escherichia coli (strain O157:H7, substra
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Aug-2002
C:Accession: B86117
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: B86117
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-75 <STO>
A:Cross-references: GB:AE005174; NID:gl2519193; PIDN:AGC59398.1; GSPDB:GN00145; UWGP:Z58
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: rpsR
C:Superfamily: Escherichia coli ribosomal protein S18

Query Match          1.3%; Score 7; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 YFRRRK 52
   |||||
Db 4 YFRRRK 10

```

```

RESULT 11
B91276
30S ribosomal subunit protein S18 [imported] - Escherichia coli (strain O157:H7, substra
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 02-Aug-2002
C:Accession: B91276
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B91276
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-75 <HAY>
A:Cross-references: GB:BA000007; PIDN:BAB38601.1; PID:gl13364655; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC85178
C:Superfamily: Escherichia coli ribosomal protein S18

Query Match          1.3%; Score 7; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 YFRRRK 52
   |||||
Db 4 YFRRRK 10

RESULT 12
A11052
30S ribosomal chain protein S18 [imported] - Salmonella enterica subsp. enterica serovar
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: A11052
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: A11052
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-75 <PAR>
A:Cross-references: GB:ALU513382; PIDN:CAD06870.1; PID:gl6505518; GSPDB:GN00176
C:Genetics:
A:Gene: rpsR
C:Superfamily: Escherichia coli ribosomal protein S18

Query Match          1.3%; Score 7; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 YFRRRK 52
   |||||
Db 4 YFRRRK 10

RESULT 13
T49763
probable magnaporin protein [imported] - Neurospora crassa
N:Alternate names: protein B24B19.350
C:Species: Neurospora crassa
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 03-Nov-2000
C:Accession: T49763
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000

```

A;Reference number: Z25022
A;Accession: T49763
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-92 <SCH>
A;Cross-references: EMBL:AL356192; GSPDB:GN00116; NCSP:B24B19.350
A;Experimental source: BAC clone B24B19; strain OR74A
C;Genetics:
A;Gene: NCSP:B24B19.350
A;Map position: 6
A;Introns: 56/1
C;Superfamily: hydrophobin HPB1

Query Match 1.3%; Score 7; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 LAAPAA 33
Db 16 LAAPAA 22

RESULT 14

S11786
nolE protein - Rhizobium leguminosarum bv. phaseoli
C;Species: Rhizobium leguminosarum bv. phaseoli
C;Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 21-Jul-2000
C;Accession: S11786
R;Davis, E.O.; Johnston, A.W.B.
Mol. Microbiol. 4, 921-932, 1990
A;Title: Analysis of three nodD genes in Rhizobium leguminosarum biovar phaseoli; nodD1
A;Reference number: S11786; MUID:91014692; PMID:2215216
A;Accession: S11786
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-110 <DAV>
A;Cross-references: GB:X54214; NID:G46231; PIDN:CAA38125.1; PID:G46232
C;Genetics:
A;Gene: nolE
C;Superfamily: Rhizobium nolE protein

Query Match 1.3%; Score 7; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 AGPSLAA 29
Db 20 AGPSLAA 26

RESULT 15

B75456
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
C;Accession: B75456
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: B75456
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-116 <WHI>
A;Cross-references: GB:AE001947; GB:AE000513; NID:G6458665; PIDN:AAF10529.1; PID:G645867
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0947
A;Map position: 1
C;Superfamily: Deinococcus radiodurans hypothetical protein DR0947

Query Match 1.3%; Score 7; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 468 HVEQARA 474
Db 8 HVEQARA 14

RESULT 16

C72654
hypothetical protein APE0663 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C;Accession: C72654
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Takah
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: C72654
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-116 <KAW>
A;Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BAA79635.1; PID:G1043421; PID:G510
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0663
C;Superfamily: Aeropyrum pernix hypothetical protein APE0663

Query Match 1.3%; Score 7; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 177 SGRFVRL 183
Db 20 SGRFVRL 26

RESULT 17

D70105
hypothetical protein BB0044 - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 28-Jul-2000
C;Accession: D70105
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943; PMID:9403685
A;Accession: D70105
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-133 <KLE>
A;Cross-references: GB:AE001118; GB:AE000783; NID:G2687921; PIDN:AAC66442.1; PID:G268793
A;Experimental source: strain B31
C;Superfamily: Borrelia burgdorferi hypothetical protein BB0044

Query Match 1.3%; Score 7; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 TOMLEKS 66
Db 31 TOMLEKS 37

RESULT 18

S75420
ribosomal protein S9 [similarity] - Sulfolobus solfataricus
N;Alternate names: protein c05003

C:Species: *Sulfolobus solfataricus*
C:Date: 11-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 03-May-2002
C:Accession: S75420
R:Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A:Title: Organizational characteristics and information content of an archaeal genome: 1
A:Reference number: S73076; MUID:97055432; PMID:8899719
A:Accession: S75420
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-140 <SEN>
A:Cross-references: EMBL:Y08257; NID:gl707772; PIDN:CAA69534.1; PID:gl707826
A:Experimental source: strain P2
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C:Genetics:
A:Start codon: GTG
C:Superfamily: Escherichia coli ribosomal protein S9
C:Keywords: protein biosynthesis; ribosome

Query Match 1.3%; Score 7; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 MEPLLLA 43
|||||
Db 53 MEPLLLA 59

RESULT 19
C70923
hypothetical protein Rv2872 - Mycobacterium tuberculosis (strain H37Rv)
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: C70923
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70923
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-147 <COL>
A:Cross-references: GB:Z74024; GB:AL123456; NID:g3250700; PIDN:CAA98372.1; PID:e248879;
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: Rv2872

Query Match 1.3%; Score 7; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 LAAPAA 33
|||||
Db 74 LAAPAA 80

RESULT 20
G72640
hypothetical protein APE0559 - *Aeropyrum pernix* (strain K1)
C:Species: *Aeropyrum pernix*
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: G72640
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyr*
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: G72640
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-147 <KAW>
A:Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA79527.1; PID:d1043313; PID:g510
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0559

Query Match 1.3%; Score 7; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 PITSSSG 178
|||||
Db 98 PITSSSG 104

RESULT 21
G84131
hypothetical protein BH3855 [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G84131
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G84131
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-151 <STO>
A:Cross-references: GB:AP001520; GB:BA000004; NID:gi0176401; PIDN:BA07574.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3855

Query Match 1.3%; Score 7; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 DTFLYLA 287
|||||
Db 48 DTFLYLA 54

RESULT 22
S73113
hypothetical protein 174 - red alga (*Porphyra purpurea*) chloroplast
C:Species: *Chloroplast Porphyra purpurea*
C:Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 10-Sep-1997
C:Accession: S73113
R:Reith, M.; Munnholland, J.
Plant Mol. Biol. Rep. 13, 333-335, 1995
A:Title: Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome.
A:Accession: S73113
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-174 <REI>
A:Cross-references: EMBL:U38804; NID:gl276652; PID:gl276658
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C:Genetics:
A:Gene: chloroplast
C:Keywords: chloroplast

Query Match 1.3%; Score 7; DB 2; Length 174;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 526 LIKQLRQ 532
|||||
Db 130 LIKQLRQ 136

RESULT 23

T16186
hypochemical protein F27D9.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16186
R:Bentley, D.
A:Description: The sequence of C. elegans cosmid F27D9.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z18473
A:Accession: T16186
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-183 <BEN>
A:Cross-references: EMBL:U49829; NID:gl203924; PID:gl203926; PIDN:AAA93383.1; CESP:F27D9
A:Gene: CESP:F27D9.4
A:Introns: 121/3

Query Match 1.3%; Score 7; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 408 ALSLAEN 414
|||||
Db 99 ALSLAEN 105

RESULT 24
E83270
hypochemical protein PA3003 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83270
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: E83270
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-209 <STO>
A:Cross-references: GB:AE004725; GB:AE004091; NID:G9949100; PIDN:AAG06391.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3003

Query Match 1.3%; Score 7; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 SLAAPAA 32
|||||
Db 12 SLAAPAA 18

RESULT 25
C64146
hypochemical protein HI0259 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
C:Accession: C64146
R:Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: C64146
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-220 <TIGR>
A:Cross-references: GB:L42023; TIGR:HI0259; GB:U32711; NID:gi573214; PIDN:AAC21924.1; PI
Query Match 1.3%; Score 7; DB 2; Length 220;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 201 RLNLTKY 207
|||||
Db 123 RLNLTKY 129

RESULT 26
JE0171
ribonuclease T2 (EC 3.1.27.1) - Japanese flying squid
N:Alternate names: ribonuclease TP
C:Species: Todarodes pacificus (Japanese flying squid)
C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 07-May-1999
C:Accession: JE0171; JN0069
R:Iwama, M.; Kusano, A.; Ogawa, Y.; Ohgi, K.; Irie, M.
Biol. Pharm. Bull. 21, 634-637, 1998
A:Title: A protease sensitive region of plant and animal ribonucleases belonging to the
A:Reference number: JE0171; MUID:98321272; PMID:9657053
A:Accession: JE0171
A:Molecule type: protein
A:Residues: 1-221 <IWA>
A:Experimental source: liver
R:Kusano, A.; Iwama, M.; Ohgi, K.; Irie, M.
Biosci. Biotechnol. Biochem. 62, 87-94, 1998
A:Title: Primary structure of a squid acid and base non-specific ribonuclease.
A:Reference number: JN0069; MUID:98162142; PMID:9501521
A:Accession: JN0069
A:Molecule type: protein
A:Residues: 1-221 <KUS>
A:Experimental source: liver
C:Comment: This protein is a base non-specific ribonuclease with guanylic acid preference
C:Superfamily: Enterobacter ribonuclease
C:Keywords: hydrolase
F:19-29, 57-102, 166-206, 183-193/Disulfide bonds: #status predicted

Query Match 1.3%; Score 7; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 298 TALNLFX 304
|||||
Db 118 TALNLFX 124

RESULT 27
T37032
probable integral membrane protein SCJ12.13c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37032
R:Murphy, L.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21619
A:Accession: T37032
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-223 <MUR>
A:Cross-references: EMBL:AL109989; PIDN:CAB53425.1; GSPDB:GN00070; SCOEDB:SCJ12.13c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCJ12.13c

Query Match 1.3%; Score 7; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 425 LGHVAVG 431
|||||

RESULT 33
AD1835
hypothetical protein alr0228 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AD1835
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuri, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD1835
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA077752.1; PID:gl7135206; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr0228

Query Match 1.3%; Score 7; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 411 LAEEEE 417
|||||

Db 33 LAEEEE 39
|||||

RESULT 34
JQ1802
B8R 31K protein precursor - vaccinia virus (strain WR)
C:Species: vaccinia virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999
C:Accession: JQ1802; F38550
R:Smith, G.L.; Chan, Y.S.; Howard, S.T.
J. Gen. Virol. 72, 1349-1376, 1991
A:Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right in
A:Reference number: JQ1767; MUID:91259063; PMID:2045793
A:Accession: JQ1802
A:Molecule type: DNA
A:Residues: 1-272 <SMI>
A:Cross-references: DBJ:D11079; NID:g222717; PIDN:BAA01838.1; PID:d1002314; PID:g222753
R:Howard, S.T.; Chan, Y.S.; Smith, G.L.
Virology 180, 633-647, 1991
A:Title: Vaccinia virus homologues of the Shope fibroma virus inverted terminal repeat P
A:Reference number: A38550; MUID:91111982; PMID:1846491
A:Accession: F38550
A:Molecule type: DNA
A:Residues: 1-272 <HOW>
A:Cross-references: GB:M58052
C:Keywords: glycoprotein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-272/Product: B8R protein #status predicted <MAT>
F:42,267/Binding site: carbohydrate (Asn) #status predicted

Query Match 1.3%; Score 7; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 202 LNLTKYS 208
|||||

Db 266 LNLTKYS 272
|||||

RESULT 35
G42526
B8R protein - vaccinia virus (strain Copenhagen)
C:Species: vaccinia virus
A:Note: host Homo sapiens (man)
C:Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 08-Apr-1994
C:Accession: G42526

R:Johnson, G.P.
submitted to GenBank, June 1990
A:Reference number: A33172
A:Accession: G42526
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-272 <JOH>

Query Match 1.3%; Score 7; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 202 LNLTKYS 208
|||||

Db 266 LNLTKYS 272
|||||

RESULT 36
T51714
probable formamidopyrimidine-DNA glycosylase 2 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000
C:Accession: T51714
R:Murphy, T.M.; Gao, M.J.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z25434
A:Accession: T51714
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-274 <MUR>
A:Cross-references: EMBL:AF099971; PIDN:AAC97953.1
C:Genetics:
A:Gene: fpg2
C:Superfamily: formamidopyrimidine-DNA glycosylase

Query Match 1.3%; Score 7; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 476 LOTASSL 482
|||||

Db 201 LOTASSL 207
|||||

RESULT 37
S69525
hypothetical protein 19 - phage HP1
C:Species: phage HP1
C:Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
C:Accession: S69525
R:Esposito, D.; Fitzmaurice, W.P.; Benjamin, R.C.; Goodman, S.D.; Waldman, A.S.; Scoocca, Nucleic Acids Res. 24, 2360-2368, 1996
A:Title: The complete nucleotide sequence of bacteriophage HP1 DNA.
A:Reference number: S69503; MUID:96279738; PMID:8710508
A:Accession: S69525
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-281 <ESP>
A:Cross-references: EMBL:U24159; NID:gl046235; PIDN:AAB09204.1; PID:gl046246
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995

Query Match 1.3%; Score 7; DB 2; Length 281;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 407 RALSALAE 413
|||||

Db 105 RALSALAE 111
|||||

RESULT 38
D72679
hypothetical protein APE0856 - Aeropyrum pernix (strain KI)

C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: D72679
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: D72679
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-282 <KAW>
A:Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BAA79836.1; PID:G5104521
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0856
C:Superfamily: Aeropyrum pernix hypothetical protein APE0856

Query Match 1.3%; Score 7; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 SLAAPAA 32
|||||
Db 22 SLAAPAA 28
|||||

RESULT 39
AH3387
hypothetical cytosolic protein BME11086 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AH3387
R:DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, M.; Mazur, M.; Goltman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AH3387
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-283 <KUR>
A:Cross-references: GB:AE008917; PIDN:RALE2267.1; PID:G17983055; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11086
A:Map position: 1

Query Match 1.3%; Score 7; DB 2; Length 283;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 469 VEQARAL 475
|||||
Db 76 VEQARAL 82
|||||

RESULT 40
CB3237
probable transcription regulator PA3269 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 20-Jan-2003
C:Accession: CB3237
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, S.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: CB3237
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-284 <STO>

C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: D72679
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: D72679
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-282 <KAW>
A:Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BAA79836.1; PID:G5104521
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0856
C:Superfamily: Aeropyrum pernix hypothetical protein APE0856

Query Match 1.3%; Score 7; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 SLAAPAA 32
|||||
Db 22 SLAAPAA 28
|||||

RESULT 39
AH3387
hypothetical cytosolic protein BME11086 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AH3387
R:DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, M.; Mazur, M.; Goltman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AH3387
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-283 <KUR>
A:Cross-references: GB:AE008917; PIDN:RALE2267.1; PID:G17983055; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11086
A:Map position: 1

Query Match 1.3%; Score 7; DB 2; Length 283;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 469 VEQARAL 475
|||||
Db 76 VEQARAL 82
|||||

RESULT 40
CB3237
probable transcription regulator PA3269 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 20-Jan-2003
C:Accession: CB3237
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, S.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: CB3237
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-284 <STO>

A:Cross-references: GB:AE004749; GB:AE004091; NID:G9949388; PIDN:AAG06657.1; GSPDB:GN001
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA3269
C:Superfamily: arabinose operon regulatory protein

Query Match 1.3%; Score 7; DB 2; Length 284;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 470 EQARALL 476
|||||
Db 229 EQARALL 235
|||||

RESULT 41
YOEC88
fimbrial adhesin K88ab precursor - Escherichia coli plasmid
N:Alternate names: K88 adhesion antigen; K88 pilin
C:Species: Escherichia coli
C:Date: 14-Nov-1983 #sequence_revision 01-Mar-1996 #text_change 16-Jul-1999
A:Accession: S07208; A03494; C45725; D45725; I41316; I41317; S20036; A43634; S24808; S24
R:Gaastra, W.; Mooi, F.R.; Stuitje, A.R.; de Graaf, F.K.
FEMS Microbiol. Lett. 12, 41-46, 1981
A:Title: The nucleotide sequence of the gene encoding the K88ab protein subunit of porci
A:Reference number: S07208
A:Accession: S07208
A:Molecule type: DNA
A:Residues: 1-285 <GAA>
A:Cross-references: EMBL:V00292; NID:G41846; PIDN:CAA23567.1; PID:G41847
A:Experimental source: variant ab
R:Kleum, P.
Eur. J. Biochem. 117, 617-627, 1981
A:Title: The complete amino-acid sequence of the K88 antigen, a fimbrial protein from Es
A:Reference number: A03494; MUID:82027186; PMID:7026236
A:Accession: A03494
A:Molecule type: protein
A:Residues: 22-285 <KLE>
A:Experimental source: strain D1721
A:Note: the K88 fimbria, 0.1-1 micrometer in length and 7 nanometers in diameter, is com
8ab variant is shown
R:Bakker, D.; Willems, P.T.; Willems, R.H.; Huismann, T.T.; Mooi, F.R.; Oudega, B.; Ste
J. Bacteriol. 174, 6350-6358, 1992
A:Title: Identification of minor fimbrial subunits involved in biosynthesis of K88 fimbri
A:Reference number: A45725; MUID:93015683; PMID:1400188
A:Contents: K12
A:Accession: C45725
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-51 <BAK>
A:Note: this sequence (subunit feaG) may be one segment of several
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:115475, NCBIP:115486)
A:Accession: D45725
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 272-285 <BA2>
A:Cross-references: EMBL:Z11710; NID:G41399; PIDN:CAA77770.1; PID:G41400
A:Note: sequence extracted from NCBI backbone (NCBIP:115490)
R:Dykes, C.W.; Halliday, I.J.; Read, M.J.; Hobden, A.N.; Harford, S.
Infect. Immun. 50, 279-283, 1985
A:Title: Nucleotide sequences of four variants of the K88 gene of porcine enterotoxigeni
A:Reference number: I41316; MUID:86007039; PMID:2412961
A:Accession: I41316
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-285 <RES>
A:Cross-references: GB:M29374; NID:G146518; PIDN:AAA24032.1; PID:G146519
A:Accession: I41317
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-94, 'T', '96-102, 'H', '104-114, 'V', '116-118, 'P', '120-125, 'K', '126-154, 'L', '156, 'R',
A:Cross-references: GB:M29375; NID:G146520; PIDN:AAA24033.1; PID:G146521

F:274/Binding site: palmitate (Cys) (covalent) #status experimental

Query Match 1.3%; Score 7; DB 1; Length 292;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 QVPRPGT 114
|||||
DB 8 QVPRPGT 14

RESULT 45

T45203

probable transport protein malg [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C:Accession: T45203

R:Robison, K.

submitted to the EMBL Data Library, September 1994

A:Reference number: Z16911

A:Accession: T45203

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-296 <KEI>

A:Cross-references: EMBL:U15180; PIDN:AAA62914.1

C:Superfamily: maltose transport protein malg

Query Match 1.3%; Score 7; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GLTTTVI 8
|||||
DB 95 GLTTTVI 101

RESULT 46

B69903

aromatic metabolite ABC transporter yodE - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C:Accession: B69903

R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bester

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen

ierch, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsira, P.; Togioni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zamstein, E.; Yoshikawa, H.; Danchin, A. Bacillus subtilis.

A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: B69903

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-303 <KUN>

A:Cross-references: GB:Z99114; GB:AL009126; NID:G2634230; PIDN:CABL3848.1; PID:G2634349

A:Experimental source: strain 168

C:Genetics:

A:Gene: yodE

C:Superfamily: Bacillus subtilis probable ABC transporter ydfo

Query Match 1.3%; Score 7; DB 1; Length 303;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 GTSLKLP 119
|||||

Db 280 GTSLKLP 286
|||||

RESULT 47

S72472

type II site-specific deoxyribonuclease (EC 3.1.21.4) Avai - Anabaena variabilis

C:Species: Anabaena variabilis

C>Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 17-Mar-1999

C:Accession: S72472

R:Ruan, H.; Lumen, K.D.; Scott, M.E.; Moran, L.S.; Moran, B.E.; Pelletier, J.J.; Hess,

Mol. Gen. Genet. 252, 695-699, 1996

A>Title: Cloning and sequence comparison of Avai and BsoBI restriction-modification syst

A:Reference number: S72471; MUID:97074885; PMID:8917312

A:Accession: S72472

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-315 <RUA>

A:Cross-references: EMBL:X98339

A:Experimental source: strain ATCC 27892

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: avair

A:Start codon: GTG

A:Description: cleaves 5'CYCGRG3' between first and second base

C:Keywords: hydrolase; restriction modification system

Query Match 1.3%; Score 7; DB 2; Length 315;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 VEQARAL 475
|||||
DB 39 VEQARAL 45

RESULT 48

AH2259

type II site-specific deoxyribonuclease [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AH2259

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi

Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AH2259

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-315 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA075330.1; PID:gl7132764; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: avair

Query Match 1.3%; Score 7; DB 2; Length 315;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 VEQARAL 475
|||||
DB 39 VEQARAL 45

RESULT 49

H82958

homoserine kinase PAS495 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: H82958

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A>Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen
 A;Reference number: A82950; MUID:20437337; PMID:10994043
 A;Accession: H82958
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-316 <STO>
 A;Cross-references: GB:AE004962; GB:AE004091; NID:g9951826; PIDN:AAG08880.1; GSPDB:GN00141
 A;Experimental source: strain PAO1
 C;Genetics:
 A;Gene: thrB; PA5495
 C;Superfamily: *Pseudomonas* homoserine kinase

Query Match 1.3%; Score 7; DB 2; Length 316;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 470 EQARALL 476
 |||||
 Db 160 EQARALL 166

RESULT 50
 B82084
 probable cobalamin biosynthesis protein CbiB VC2380 [imported] - *Vibrio cholerae* (strain 569B)
 C;Species: *Vibrio cholerae*
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 23-Sep-2002
 A;Accession: B82084
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A;Reference number: A82035; MUID:20406833; PMID:10952301
 A;Accession: B82084
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-317 <HEI>
 A;Cross-references: GB:AE004308; GB:AE003852; NID:g9656946; PIDN:AAF95523.1; GSPDB:GN00141
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor
 C;Genetics:
 A;Gene: VC2380
 A;Map position: 1
 C;Superfamily: cobalamin biosynthesis protein D

Query Match 1.3%; Score 7; DB 2; Length 317;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 QARALLQ 477
 |||||
 Db 123 QARALLQ 129

RESULT 51
 D69110
 homoserine O-acetyltransferase - *Methanobacterium thermoautotrophicum* (strain Delta H)
 C;Species: *Methanobacterium thermoautotrophicum*
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 23-Dec-2002
 A;Accession: D69110
 R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.; Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A>Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: function
 A;Reference number: A69000; MUID:98037514; PMID:9371463
 A;Accession: D69110
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-319 <MTH>

A;Cross-references: GB:AE000935; GB:AE000666; NID:g2622945; PIDN:AAB86286.1; PID:g262295
 A;Experimental source: strain Delta H
 C;Genetics:
 A;Gene: MTH1820
 C;Superfamily: homoserine acetyltransferase

Query Match 1.3%; Score 7; DB 2; Length 319;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 406 ERALSLSA 412
 |||||
 Db 191 ERALSLSA 197

RESULT 52
 G87494
 hypothetical protein CC1980 [imported] - *Caulobacter crescentus*
 C;Species: *Caulobacter crescentus*
 C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C;Accession: G87494
 R;Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A>Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A;Reference number: A87249; MUID:21173698; PMID:11259647
 A;Accession: G87494
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-331 <STO>
 A;Cross-references: GB:AE005673; NID:gl3423445; PIDN:AAK23955.1; GSPDB:GN00148
 C;Genetics:
 A;Gene: CC1980
 C;Superfamily: conserved hypothetical protein H70176

Query Match 1.3%; Score 7; DB 2; Length 331;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 GVDPRPL 19
 |||||
 Db 131 GVDPRPL 137

RESULT 53
 D96515
 hypothetical protein F16N3.22 [imported] - *Arabidopsis thaliana*
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: D96515
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: D96515
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-331 <STO>
 A;Cross-references: GB:AE005173; NID:g5668810; PIDN:AAD46036.1; GSPDB:GN00141
 C;Genetics:
 A;Gene: F16N3.22
 A;Map position: 1

Query Match 1.3%; Score 7; DB 2; Length 331;

Best Local Similarity 100.0%; Pred. No. 1e+02; Mismatches 0; Gaps 0;
 Matches 7; Conservative 0; Indels 0;
 QY 475 LLOQTASS 481
 DB 97 LLOQTASS 103
 |||||
 |||||

RESULT 54
 A:2837
 C:Species: Agrobacterium tumefaciens (strain C58, Cere)
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AD2837
 R:Wood, D.W.; Scubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; PMID:21608550; PMID:11743193
 A:Accession: AD2837
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-337 <KUR>
 A:Cross-references: GB:AE008688; PIDN:AAL43114.1; PID:gl17740586; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu2124
 A:Map position: circular chromosome

Query Match 1.3%; Score 7; DB 2; Length 337;
 Best Local Similarity 100.0%; Pred. No. 1e+02; Mismatches 0; Gaps 0;
 Matches 7; Conservative 0; Indels 0;
 QY 362 DOPEIAL 368
 DB 232 DOPEIAL 238
 |||||
 |||||

RESULT 55
 F97614
 Hypothetical protein AGR_C_3852 [imported] - Agrobacterium tumefaciens (strain C58, Cere)
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
 C:Accession: F97614
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
 A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359; PMID:21608551; PMID:11743194
 A:Accession: F97614
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-337 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK87871.1; PID:gl15157259; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C_3852
 A:Map position: circular chromosome

Query Match 1.3%; Score 7; DB 2; Length 337;
 Best Local Similarity 100.0%; Pred. No. 1e+02; Mismatches 0; Gaps 0;
 Matches 7; Conservative 0; Indels 0;
 QY 362 DOPEIAL 368
 DB 232 DOPEIAL 238
 |||||
 |||||

RESULT 56
 LUPY1
 annexin I type 1 - pigeon

N:Alternate names: calpactin II; calpactin/lipocortin homolog; cropsac 35K protein; phos
 C:Species: Columba livia (domestic pigeon)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Feb-1997
 C:Accession: A40153; PH0846
 R:Horseman, N.D.
 Mol. Endocrinol. 3, 773-779, 1989
 A:Title: A prolactin-inducible gene product which is a member of the calpactin/lipocortin
 A:Reference number: A40153; PMID:89330493; PMID:2526923
 A:Accession: A40153
 A:Molecule type: mRNA
 A:Residues: 1-341 <HOR>
 A:Cross-references: GB:M22635
 R:Hitti, Y.S.; Horseman, N.D.
 Gene 103, 185-192, 1991
 A:Title: Structure of the gene encoding columbid annexin I cp35.
 A:Reference number: PH0846; PMID:91365244; PMID:1832409
 A:Accession: PH0846
 A:Molecule type: DNA
 A:Residues: 1-53 <HIT>
 C:Genetics:
 A:Gene: cp35
 A:Introns: 22/3; 54/1
 C:Superfamily: annexin I; annexin repeat homology
 C:Keywords: calcium; duplication; endonexin fold; inflammation; phospholipase A2 inhibit
 F:2-340/Product: annexin I #status predicted <MAT>
 F:40-111/Domain: annexin repeat homology <AX1>
 F:51-67/Region: endonexin fold #status predicted
 F:112-183/Domain: annexin repeat homology <AX2>
 F:123-139/Region: endonexin fold #status predicted
 F:195-266/Domain: annexin repeat homology <AX3>
 F:207-223/Region: endonexin fold #status predicted
 F:270-341/Domain: annexin repeat homology <AX4>
 F:281-297/Region: endonexin fold #status predicted
 F:211/Binding site: phosphate (thr) (covalent) (by CAMP-dependent kinase) #status predic
 Query Match 1.3%; Score 7; DB 1; Length 341;
 Best Local Similarity 100.0%; Pred. No. 1e+02; Mismatches 0; Gaps 0;
 Matches 7; Conservative 0; Indels 0;
 QY 336 YYKEVLK 342
 DB 150 YYKEVLK 156
 |||||
 |||||

RESULT 57
 A4118
 annexin I type 2 - pigeon
 N:Alternate names: calpactin II; calpactin/lipocortin homolog; cp37 protein; cropsac 37K
 C:Species: Columba livia (domestic pigeon)
 C:Date: 31-Dec-1993 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: A44118; B44118; A38909; I51112
 R:Haigler, H.T.; Mangili, J.A.; Gao, Y.; Jones, J.; Horseman, N.D.
 J. Biol. Chem. 267, 19123-19129, 1992
 A:Title: Identification and characterization of columbid annexin Icp37. Insights into th
 A:Reference number: A44118; PMID:92406850; PMID:1388165
 A:Accession: A44118
 A:Molecule type: mRNA
 A:Residues: 1-343 <HA1>
 A:Note: sequence extracted from NCBI backbone (NCBIN:113807, NCBI:P:113808)
 A:Accession: B44118
 A:Molecule type: protein
 A:Residues: 13-15 'X', 17, 'X', 19-28 'X', 30-50 <HA2>
 R:Haigler, H.T.; Mangili, J.A.; Gao, Y.; Jones, J.; Horseman, N.D.
 submitted to GenBank, October 1992
 A:Reference number: A38909
 A:Accession: A38909
 A:Molecule type: mRNA
 A:Residues: 1-73 'HR', 76-343 <HA3>
 A:Cross-references: GB:M91008; NID:G213521; PIDN:AAA49447.1; PID:G213522
 A:Note: amino end is blocked in the mature form
 R:Gao, Y.; Horseman, N.D.
 Gene 143, 179-186, 1994
 A:Title: Structural and functional divergences of the columbid annexin I-encoding cp37

A:Reference number: I51112; MUID:94266150; PMID:8206371
 A:Accession: I51112
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-154 <GAO>
 A:Cross-references: GB:L02504; NID:g213535; PIDN:AAA20674.1; PID:g529725
 C:Genetics:
 A:Gene: cp37
 A:Introns: 22/3; 55/1; 86/3; 124/3
 C:Superfamily: annexin I; annexin repeat homology
 C:Keywords: acetylated amino end; calcium binding; duplication; endonexin fold; inflammation
 F:2-342/Product: annexin I #status predicted <MAT>
 F:41-112/Domain: annexin repeat homology <AX1>
 F:52-68/Region: endonexin fold #status predicted
 F:113-184/Domain: annexin repeat homology <AX2>
 F:114-140/Region: endonexin fold #status predicted
 F:196-268/Domain: annexin repeat homology <AX3>
 F:208-224/Region: endonexin fold #status predicted
 F:272-343/Domain: annexin repeat homology <AX4>
 F:283-299/Region: endonexin fold #status predicted
 F:21/Modified site: acetylated amino end (Ala) (in mature form) #status predicted
 F:24/Binding site: phosphate (Tyr) (covalent) (by EGF receptor/kinase) #status predicted
 F:24/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 1.3%; Score 7; DB 1; Length 343;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 336 YKKEVLK 342
 |||||
 Db 151 YKKEVLK 157

RESULT 58
 G82690
 aspartate-B-semialdehyde dehydrogenase XF1371 [imported] - Xylella fastidiosa (strain 9a)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000
 C:Accession: G82690
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequences
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: G82690
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-344 <SIM>
 A:Cross-references: GB:AE003968; NID:g9106363; PIDN:AAF84180.1; GSPDB:GN001
 A:Experimental source: strain 9a5C
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H. as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF1371
 C:Superfamily: aspartate-semialdehyde dehydrogenase

Query Match 1.3%; Score 7; DB 2; Length 344;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 470 EQARALL 476

Db 264 EQARALL 270
 |||||
 RESULT 59
 F71372
 probable protoporphyrinogen oxidase (hemK) - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
 C:Accession: F71372
 R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin rson, J.; Khailak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A:Reference number: A71250; MUID:98332770; PMID:9665876
 A:Accession: F71372
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-348 <COL>
 A:Cross-references: GB:AE001190; GB:AE000520; NID:g3322303; PIDN:AA65048.1; PID:g332231 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0052

Query Match 1.3%; Score 7; DB 2; Length 348;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 QARALLQ 477
 |||||
 Db 252 QARALLQ 258

RESULT 60
 A98332
 malate dehydrogenase PAB1791 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
 C:Species: Agrobacterium tumefaciens
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
 C:Accession: A98332
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wolam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum. A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: A98332
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-358 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK90179.1; PID:g15160184; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L_3209
 A:Map position: linear chromosome
 C:Superfamily: malate dehydrogenase ylbC

Query Match 1.3%; Score 7; DB 2; Length 358;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 GTRLGVD 15
 |||||
 Db 45 GTRLGVD 51

RESULT 61
 S17905
 hypothetical protein M1 - La France disease virus
 C:Species: La France disease virus
 C:Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 21-Jul-2000
 C:Accession: S17905
 R:Harmson, M.C.; Tolner, B.; Kram, A.; Go, S.J.; de Haan, A.; Wessels, J.G.H. Curr. Genet. 20, 137-144, 1991
 A:Title: Sequences of three dsRNAs associated with La France disease of the cultivated m

A:Reference number: S17904; MUID:92035060; PMID:1934110
A:Accession: S17905
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-359 <HAR>
A:Cross-references: GB:D10829; GB:D00483; NID:g221977; PIDN:BAA01611.1; PID:g221978
C:Superfamily: La France disease virus hypothetical protein M1

Query Match 1.3%; Score 7; DB 2; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 WLALSTE 240
|||||
Db 214 WLALSTE 220

RESULT 62
D86327
protein F18014.17 [imported] - Arabidopsis thaliana.
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
A:Accession: D86327
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huiziar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Cazzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
Rizzo, M.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, A.; Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: D86327
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-361 <STO>
A:Cross-references: GB:AE005172; NID:98778439; PIDN:AAF79447.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 1.3%; Score 7; DB 2; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 215 KACLSTIS 221
|||||
Db 75 KACLSTIS 81

RESULT 63
B75132
cell division control protein 48, aaa family (cdc48-2) PAB1789 - Pyrococcus abyssi (stra
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
A:Accession: B75132
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: B75132
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-363 <KAW>
A:Cross-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CAB49771.1; PID:g545828
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1789
C:Superfamily: ATP-dependent 26S proteinase; FtsH/SEC18/CDC48-type ATP-binding domain hc

Query Match 1.3%; Score 7; DB 2; Length 363;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 405 FERALSL 411
|||||
Db 16 FERALSL 22

RESULT 64
C87684
conserved hypothetical protein CC3509 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
A:Accession: C87684
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: C87684
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-365 <STO>
A:Cross-references: GB:AE005673; NID:g13425237; PIDN:AAK25471.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3509

Query Match 1.3%; Score 7; DB 2; Length 365;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 AAPAAMS 34
|||||
Db 318 AAPAAMS 324

RESULT 65
B69072
hypothetical protein MTH1540 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 04-Mar-2000
A:Accession: B69072
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.
Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: B69072
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-367 <MTH>
A:Cross-references: GB:AE000914; GB:AE000666; NID:g2622656; PIDN:AAB86014.1; PID:g262266
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1540
A:Start codon: TTG
C:Superfamily: Methanobacterium thermoautotrophicum hypothetical protein MTH1540

Query Match 1.3%; Score 7; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 411 LAENEE 417
|||||
Db 201 LAENEE 207

RESULT 66
A71025
hypothetical protein PH1494 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii

C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C;Accession: A71025
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yanazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Accession number: A71000; MUID:98344137; PMID:9679194
A;Accession: A71025
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-369 <KAW>
A;Cross-references: GB:AP000006; NID:G3236133; PIDN:BAA30601.1; PID:G3257918
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1494
C;Superfamily: Pyrococcus horikoshii hypothetical protein PH1494

Query Match 1.3%; Score 7; DB 2; Length 369;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 200 SRLNLTK 206
Db 121 SRLNLTK 127
|||||

RESULT 67
C81263
probable integral membrane protein Cj1662 [imported] - Campylobacter jejuni (strain NCTC
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C;Accession: C81263
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A;Reference number: A81250; MUID:20150912; PMID:10698204
A;Accession: C81263
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-372 <PAR>
A;Cross-references: GB:AL139079; GB:AL111168; NID:G6968971; PIDN:CAB73649.1; PID:G696907
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: Cj1662

Query Match 1.3%; Score 7; DB 2; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 LKARALT 81
Db 83 LKARALT 89
|||||

RESULT 68
AD1256
conserved hypothetical protein lmo1452 [imported] - Listeria monocytogenes (strain EGD-e
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AD1256
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AD1256
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-373 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99530.1; PID:G16410881; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo1452

Query Match 1.3%; Score 7; DB 2; Length 373;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 443 RLALVNN 449
Db 152 RLALVNN 158
|||||

RESULT 69
AH1618
conserved hypothetical protein lin1489 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AH1618
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1618
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-373 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC96720.1; PID:G16413962; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin1489

Query Match 1.3%; Score 7; DB 2; Length 373;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 443 RLALVNN 449
Db 152 RLALVNN 158
|||||

RESULT 70
S48639
fructose-bisphosphate aldolase (EC 4.1.2.13) precursor - Chlamydomonas reinhardtii
N;Alternate names: fructose-1,6-bisphosphate aldolase
C;Species: Chlamydomonas reinhardtii
C;Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 22-Jun-1999
C;Accession: S48639; S58485; S58486; S34367
R;Schnarrenberger, C.; Pelzer-Reith, B.; Yatsuki, H.; Freund, S.; Jacobshagen, S.; Hori,
Arch. Biochem. Biophys. 313, 173-178, 1994
A;Title: Expression and sequence of the only detectable aldolase in Chlamydomonas reinha
A;Reference number: S48639; MUID:94330714; PMID:8053679
A;Accession: S48639
A;Molecule type: mRNA
A;Residues: 1-374 <SCH>
A;Cross-references: EMBL:X69969; NID:G312755; PIDN:CAA49590.1; PID:G312756
R;Pelzer-Reith, B.; Freund, S.; Schnarrenberger, C.; Yatsuki, H.; Hori, K.
Mol. Gen. Genet. 248, 481-486, 1995
A;Title: The plastid aldolase gene from Chlamydomonas reinhardtii: intron/exon organizat
A;Reference number: S58485; MUID:96004769; PMID:7565612
A;Accession: S58485
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 28-374 <PEL>
A;Cross-references: EMBL:X85495
R;Pelzer-Reith, B.
submitted to the EMBL Data Library, March 1995

A:Reference number: S58486
A:Accession: S58486
A:Molecule type: DNA
A:Residues: 1-367,370-374 <PEW>
A:Cross-references: EMBL:X85495
C:Genetics:
A:Gene: ALDCHL

A:Genome: nuclear
A:Introns: 209/3; 235/3; 325/3
C:Superfamily: fructose-bisphosphate aldolase
C:Keywords: aldehyde-lyase; carbon-carbon lyase; chloroplast
F:1-27/Domains: transit peptide (chloroplast) #status predicted <TNP>
F:28-374/Product: fructose-bisphosphate aldolase #status predicted <MAT>
F:162,243,374/Active site: Lys, Lys, Tyr #status predicted

Query Match 1.3%; Score 7; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 SAAEYK 338
|||
DB 149 SAAEYK 155
|||

RESULT 71

AG2460
hypothetical protein alr5239 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AG2460
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AG2460
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-389 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA076938.1; PID:gi7134378; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr5239

Query Match 1.3%; Score 7; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 GGPSQAV 131
|||
DB 15 GGPSQAV 21
|||

RESULT 72

E96565
hypothetical protein F6D8.28 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E96565
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huijzer, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: E96565
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-390 <STO>
A:Cross-references: GB:AE005173; NID:G5903053; PIDN:AA055612.1; GSPDB:GN00141
C:Genetics:
A:Gene: F6D8.28
A:Map position: 1

Query Match 1.3%; Score 7; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 476 LQTASSL 482
|||
DB 201 LQTASSL 207
|||

RESULT 73

T51713
probable formamidopyrimidine-DNA glycosylase 1 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: T51713
R:Murphy, T.M.; Gao, M.J.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z25434
A:Accession: T51713

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-390 <MUR>
A:Cross-references: EMBL:AF099970; PIDN:AAC97952.1
C:Genetics:
A:Gene: fpg1

Query Match 1.3%; Score 7; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 476 LQTASSL 482
|||
DB 201 LQTASSL 207
|||

RESULT 74

T35440
probable polyamine-binding lipoprotein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T35440
R:Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1998
A:Reference number: Z21578
A:Accession: T35440

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-399 <MUR>
A:Cross-references: EMBL:AL031035; PIDN:CAA19894.1; GSPDB:GN00070; SCOEDB:SC6A9.09C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC6A9.09C

Query Match 1.3%; Score 7; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 GRFVRLG 184
|||
DB 130 GRFVRLG 136
|||

RESULT 75

A70416
ATP-dependent clp proteinase (EC 3.4.21.-) regulatory chain X - Aquifex aeolicus
C:Species: Aquifex aeolicus


C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Feb-2001
 C;Accession: A70416
 R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
 V.
 Nature 392, 353-358, 1998
 A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A;Reference number: A70300; MUID:98196666; PMID:9537320
 A;Accession: A70416
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-412 <AOF>
 A;Cross-references: GB:AE000735; NID:G2983749; PIDN:AAC07316.1; PID:G2983756; GB:AE00069
 A;Experimental source: strain VF5
 C;Genetics:
 A;Gene: clpX
 C;Complex: heterodimer; proteolytic chain (clpP) and ATP-binding regulatory chain (clpX)
 C;Function:
 A;Description: ATP-dependent specificity component of the clpP proteinase
 C;Superfamily: heat shock protein hslU; FtsH/SEC18/CDC48-type ATP-binding domain homolog
 C;Keywords: ATP; heterodimer; hydrolase; molecular chaperone; nucleotide binding; P-loop
 F;10-36/Region: zinc finger CCCC motif
 F;117-124/Region: nucleotide-binding motif A (P-loop)

Query Match 1.3%; Score 7; DB 2; Length 412;
 Best Local Similarity 100.0%; Pred.No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 VYIDEID 90
 |||||
 Db 179 VYIDEID 185

Search completed: February 10, 2004, 13:29:07
 Job time : 25 secs

THIS PAGE BLANK (USPTO)

<i>Application Number</i> 	Application No. 10/380,783	Applicant(s) NAGATO ET AL.	
	Examiner John M Ford	Art Unit 1624	

